Challenges and Advances
in Computational Chemistry and Physics 14
Series Editor: J.Leszcynski

Malgorzata Baranska *Editor*

Optical Spectroscopy and Computational Methods in Biology and Medicine

Springer

Optical Spectroscopy and Computational Methods in Biology and Medicine

CHALLENGES AND ADVANCES IN COMPUTATIONAL CHEMISTRY AND PHYSICS

Volume14

Series Editor:

JERZY LESZCZYNSKI
Department of Chemistry, Jackson State University, U.S.A.

For further volumes:
http://www.springer.com/series/6918

Optical Spectroscopy and Computational Methods in Biology and Medicine

Edited by

Malgorzata Baranska
Faculty of Chemistry, Jagiellonian University, Ingardena 3, 30-060 Krakow, Poland

Springer

Editor
Malgorzata Baranska
Faculty of Chemistry
Jagiellonian University
Ingardena 3
Krakow
Poland

ISBN 978-94-007-7831-3 ISBN 978-94-007-7832-0 (eBook)
DOI 10.1007/978-94-007-7832-0
Springer Dordrecht Heidelberg New York London

Library of Congress Control Number: 2013954304

© Springer Science+Business Media Dordrecht 2014
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

This book is focus mainly on vibrational spectroscopy, Raman scattering and infrared absorption, and their applications in biology and medicine. When talking about new techniques, beside imaging or surface-enhanced spectroscopy also the investigation of optical activity of the samples became significantly common. In view of the fact that beside vibrational spectroscopy also electronic absorption is presented here as a potent and complementary tool to study absolute configuration, the overall topic of this book is named "optical spectroscopy", however the main part is devoted to vibrational methods.

Since the sample investigation is often complicated, especially when a specific analyte is spread in the biomatrix and occurs in low concentration, it is necessary to support the analysis by theoretical methods including chemometrics or 2D correlation spectroscopy. Demonstrated here approach connecting optical spectroscopy with computer methods is shown as a powerful tool that can be successfully used in biology and medicine to study very complex samples.

The book contains a short introduction followed by a set of examples presenting results obtained by using spectroscopy combined with calculations. New techniques and its special application are based on Surface-Enhanced Raman Spectroscopy (SERS), Raman Optical Activity (ROA), Vibrational Circular Dichroism (VCD), Electronic Circular Dichroism (ECD) and matrix isolation spectroscopy. For biological applications the analysis of secondary and primary plant components in the tissue and single cells is demonstrated. After that, biomedical applications of optical spectroscopy related to disease diagnosis including the identification of illnesses biomarkers. A broad qualitative, quantitative and structural study is possible when the advanced computer methods are used.

Malgorzata Baranska

Contents

Part I Introduction

1 General Overview on Vibrational Spectroscopy Applied in Biology and Medicine 3
Malgorzata Baranska, Maciej Roman and Katarzyna Majzner

2 Integrating Optical Spectroscopy and Chemometric Methods 15
Karolina Jagiełło, Anita Sosnowska, Jan Mazerski and Tomasz Puzyn

Part II New Techniques and Special Applications

3 Surface-enhanced Raman Scattering (SERS) in Bioscience: A Review of Application 29
Jolanta Bukowska and Piotr Piotrowski

4 Raman Optical Activity of Biological Samples..... 61
Katarzyna Chruszcz-Lipska and Ewan W. Blanch

5 α-Amino Acids In Water: A Review Of VCD And Roa Spectra 83
Jan Cz. Dobrowolski, Piotr F.J. Lipiński, Joanna E. Rode and Joanna Sadlej

6 Electronic Circular Dichroism Spectroscopy in Structural Analysis of Biomolecular Systems 161
Magdalena Pecul and Wojciech Dzwolak

7 Biologically Relevant Molecules Studied in Low Temperature Inert Matrices..... 179
Rui Fausto and Nihal Kuş

8 Two-Dimensional Mid-Infrared Correlation Spectroscopy in Protein Research 213
Bogusława Czarnik-Matusewicz and Young Mee Jung

Part III Biological Applications

9 Qualitative and Quantitative FT-Raman Analysis of Plants 253
Hartwig Schulz

10 Chemometric Analysis of Raman and IR Spectra of Natural Dyes 279
Anna Rygula and Paweł Miskowiec

11 Structural and Spatial Analysis of Carotenoids in a Single Cell Monitored by Raman Spectroscopy 309
Agnieszka Kaczor and Marta Pilarczyk

Part IV Biomedical Applications

12 Raman Spectra of Solid Amino Acids: Spectral Correlation Analysis as the First Step Towards Identification by Raman Spectroscopy 329
Tomasz Roliński, Sylwester Gawinkowski, Agnieszka Kamińska and Jacek Waluk

13 Vibrational Spectroscopy: Disease Diagnostics and Beyond 355
Hugh J. Byrne, Kamila M. Ostrowska, Haq Nawaz, Jennifer Dorney, Aidan D. Meade, Franck Bonnier and Fiona M. Lyng

14 Quantitative Analysis of Disease Biomarkers Using Surface-Enhanced Raman Scattering Spectroscopy 401
Hyangah Chon and Jaebum Choo

15 FTIR Imaging of Tissues: Techniques and Methods of Analysis 419
Kamilla Malek, Bayden R. Wood and Keith R. Bambery

16 Fourier Transform Infrared (Ft-Ir) Spectroscopic Imaging for Solid Tumor Histopathology 475
Sreeradha Biswas, Michael J. Walsh and Rohit Bhargava

17 Vibrational Microspectroscopy for Analysis of Atherosclerotic Arteries 505
T.P. Wróbel, A. Fedorowicz, Ł. Mateuszuk, E. Maślak, A. Jasztal, S. Chlopicki and K.M. Marzec

Index 537

List of Contributors

Keith R. Bambery Australian Synchrotron, Clayton, VIC, Australia

Malgorzata Baranska Faculty of Chemistry, Jagiellonian University, Krakow, Poland

Rohit Bhargava Departments of Bioengineering, Mechanical Science and Engineering, Electrical and Computer Engineering, Chemical and Biomolecular Engineering and Chemistry, University of Illinois Urbana-Champaign, Urbana, IL, USA

Sreeradha Biswas Biophysics program, University of Illinois Urbana-Champaign, Urbana, IL, USA

Ewan W. Blanch Manchester Institute of Biotechnology and Faculty of Life Sciences, The University of Manchester, Manchester, UK

Franck Bonnier Focas Research Institute, Dublin Institute of Technology, Dublin 8 Ireland

Jolanta Bukowska Department of Chemistry, University of Warsaw, Warszawa, Poland

Hugh J. Byrne Focas Research Institute, Dublin Institute of Technology, Dublin 8 Ireland

S. Chlopicki Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

Department of Experimental Pharmacology, Chair of Pharmacology, Jagiellonian University, Medical College, Krakow, Poland

Hyangah Chon Department of Bionano Engineering and Center for Integrated Human Sensing System, Hanyang University, Ansan, Korea

Jaebum Choo Department of Bionano Engineering and Center for Integrated Human Sensing System, Hanyang University, Ansan, Korea

Katarzyna Chruszcz-Lipska Jagiellonian Centre for Experimantal Therapeutics, Jagiellonian University, Kraków, Poland

Bogusława Czarnik-Matusewicz Faculty of Chemistry, University of Wrocław, Wrocław, Poland

Jan Cz. Dobrowolski Industrial Chemistry Research Institute, Warsaw, Poland

National Medicines Institute, Warsaw, Poland

Jennifer Dorney Focas Research Institute, Dublin Institute of Technology, Dublin 8 Ireland

Wojciech Dzwolak Faculty of Chemistry, University of Warsaw, Warszawa, Poland

Rui Fausto Department of Chemistry, University of Coimbra, Coimbra, Portugal

A. Fedorowicz Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

Sylwester Gawinkowski Institute of Physical Chemistry, Polish Academy of Sciences, Warsaw, Poland

Karolina Jagiełło Laboratory of Environmental Chemometrics, Faculty of Chemistry, University of Gdansk, Gdansk, Poland

A. Jasztal Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

Young Mee Jung Department of Chemistry, Kangwon National University, Chuncheon, Korea

Agnieszka Kaczor Faculty of Chemistry, Jagiellonian University, Krakow, Poland

Jagiellonian Centre of Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

Agnieszka Kamińska Institute of Physical Chemistry, Polish Academy of Sciences, Warsaw, Poland

Nihal Kuş Department of Chemistry, University of Coimbra, Coimbra, Portugal

Department of Physics, Anadolu University, Eskişehir, Turkey

Piotr F.J. Lipiński Mossakowski Medical Research Centre,Polish Academy of Sciences, Warsaw, Poland

Fiona M. Lyng Focas Research Institute, Dublin Institute of Technology, Dublin 8 Ireland

Katarzyna Majzner Faculty of Chemistry, Jagiellonian University, Krakow, Poland

Kamilla Malek Faculty of Chemistry and Jagiellonian Centre for Experimental Therapeutics, Jagiellonian University, Krakow, Poland

K.M. Marzec Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

E. Maślak Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

Ł. Mateuszuk Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

Jan Mazerski Department of Pharmaceuticals Technology and Biochemistry, Chemical Faculty, Gdansk University of Technology, Gdansk, Poland

Aidan D. Meade Focas Research Institute, Dublin Institute of Technology, Dublin 8 Ireland

Paweł Miskowiec Faculty of Chemistry, Jagiellonian University, Krakow, Poland

Haq Nawaz Focas Research Institute, Dublin Institute of Technology, Dublin 8 Ireland

Kamila M. Ostrowska Focas Research Institute, Dublin Institute of Technology, Dublin 8 Ireland

Magdalena Pecul Faculty of Chemistry, University of Warsaw, Warszawa, Poland

Marta Pilarczyk Faculty of Chemistry, Jagiellonian University, Krakow, Poland

Piotr Piotrowski Department of Chemistry, University of Warsaw, Warszawa, Poland

Tomasz Puzyn Laboratory of Environmental Chemometrics, Faculty of Chemistry, University of Gdansk, Gdansk, Poland

Joanna E. Rode Industrial Chemistry Research Institute, Warsaw, Poland

Tomasz Roliński Institute of Physical Chemistry, Polish Academy of Sciences, Warsaw, Poland

Maciej Roman Faculty of Chemistry, Jagiellonian University, Krakow, Poland

Anna Rygula Faculty of Chemistry, Jagiellonian University, Krakow, Poland

Joanna Sadlej National Medicines Institute, Warsaw, Poland

Faculty of Chemistry, University of Warsaw, Warsaw, Poland

Hartwig Schulz Julius Kühn-Institute, Institute for Ecological Chemistry, Plant Analysis and Stored Product Protection, Federal Research Centre for Cultivated Plants, Berlin, Germany

Anita Sosnowska Laboratory of Environmental Chemometrics, Faculty of Chemistry, University of Gdansk, Gdansk, Poland

Michael J. Walsh Beckman Institute for Advanced Science and Technology, University of Illinois Urbana-Champaign, Urbana, IL, USA

Jacek Waluk Institute of Physical Chemistry, Polish Academy of Sciences, Warsaw, Poland

Bayden R. Wood Centre for Biospectroscopy and School of Chemistry, Monash University, Victoria, Australia

T.P. Wróbel Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

Faculty of Chemistry, Jagiellonian University, Krakow, Poland

Part I
Introduction

Chapter 1
General Overview on Vibrational Spectroscopy Applied in Biology and Medicine

Malgorzata Baranska, Maciej Roman and Katarzyna Majzner

Abstract This chapter contains a short introduction to vibrational spectroscopy followed by an overview on its biological and biomedical applications. The spectroscopic techniques discussed in the book and their special advantages are briefly listed, i.e. Surface-Enhanced Raman Spectroscopy (SERS), Raman Optical Activity (ROA), Vibrational Circular Dichroism (VCD), Electronic Circular Dichroism (ECD) and matrix isolation.

The potential of vibrational spectroscopy is demonstrated by the current state of the art in secondary and primary plant components analysis performed in the tissue and from the single cells. Both Raman and IR spectroscopy are shown as powerful tools in medical diagnosis, cytology and histopathology. A brief overview on biomedical vibrational spectroscopy used to investigate lifestyle diseases is provided.

Keywords Surface-Enhanced Raman Spectroscopy (SERS) • Raman Optical Activity (ROA) • Vibrational Circular Dichroism (VCD) • Electronic Circular Dichroism (ECD) • Matrix isolation • Biological and biomedical application

1.1 Vibrational Spectroscopy Methods

Infrared and Raman spectroscopy are among the most widely used techniques in natural sciences. This is due to the fact that vibrational spectroscopy can be applied not only for samples in milligram quantities but microsampling and trace analysis is nowadays done routinely, and measurements take from minutes to fractions of seconds.

The selection rules, which formally constrains the possible transitions of a molecular system, are different for IR and Raman spectroscopy. The classical theory

M. Baranska (✉) · M. Roman · K. Majzner
Faculty of Chemistry, Jagiellonian University, Ingardena 3,
30-060, Krakow, Poland
e-mail: baranska@chemia.uj.edu.pl

M. Baranska (ed.), *Optical Spectroscopy and Computational Methods in Biology and Medicine,* Challenges and Advances in Computational Chemistry and Physics 14,
DOI 10.1007/978-94-007-7832-0_1, © Springer Science+Business Media Dordrecht 2014

has lead to the following criteria: a given normal vibration of a molecule may appear in the IR spectrum if at least one component of the electric dipole moment of the molecule changes during the vibration whereas respectively for the Raman spectrum: at least one component of the polarizability tensor changes. However according to quantum theory, oscillation of a permanent or induced dipole does not produce any radiation. For radiation to be emitted or absorbed, a transition should occur between different vibrational energy levels of the molecule which may be concomitant with loss or gain of energy in the form of electromagnetic radiations. The classical analogue of excitation of a molecule to a higher energy level is vibration with increased amplitude [1].

Also a spatial resolution of both vibrational methods is different. According to the Rayleigh criterion, which defines the resolution obtained from the diffraction limited grating spectrometers and also by a Fourier transform spectrometer that uses triangular apodization, IR spectroscopy is at the limit of 10–12 microns whereas for Raman scattering it is 0.3 microns (for the excitation of 488 nm). Various sampling technique can improve a spatial resolution, however their employment is also limited.

IR and Raman spectroscopy are complimentary methods. Among 3N-6 vibrational modes (N—the number of atoms in a molecule), for some specific functional groups a characteristic frequency can be used for their identification, quantification or conformational study.

Both methods in practice are used as individual techniques, sometimes restricted only to special applications but also providing unusual possibilities and advantages. For the purpose of this book the results of Raman and IR imaging, Surface-Enhanced Raman Spectroscopy (SERS), Raman Optical Activity (ROA), Vibrational Circular Dichroism (VCD), and matrix isolation spectroscopy are presented. Additionally to dichroism and optical activity in vibrational spectroscopy, Electronic Circular Dichroism (ECD) is introduced.

Vibrational spectroscopic imaging greatly extend the possibilities of the conventional IR and Raman spectroscopy. Spectral data can be represented as a picture, presenting chemical information simultaneously from thousands of pixels. The pixel brightness or colour can be derived from any number of spectral parameters, ranging from simple vibrational peak intensities to multivariate parameters. The ability to explore spectral information in a spatially resolve manner can give significant insight into organisational motifs and aid the prediction of molecular functional behaviour. Hopefully the usefulness of an integrated spectral and spatial techniques will result in their application to routine analysis, with a special focus on biological samples and medical diagnosis [1].

Surface Enhanced Raman Spectroscopy (SERS) is a useful analytical technique employed for the identification accompanied by investigation of the most probable orientation of molecules adsorbed on the metal surface [2]. Although SERS spectroscopy is widely applied, the mechanism leading to the surface enhancement is not completely understood yet. Three possible contributions to the enhancement factor have been identified: (i) the surface plasmons resonance in the metal nanoparticles, (ii) a charge-transfer resonance involving shift of electrons between

the molecule and the conduction band of the metal, and (iii) resonances within the molecule itself [3].

VCD, ROA and ECD are applied to study enantiomers, but beside small molecules also large and complex biopolymers e.g. proteins and DNA can be analysed. With the help of *ab initio* calculations the identification of absolute configuration of the molecule is possible. The measurements requires the use of left- and right- circularly polarized light, the spectrometers must be very sensitive, and usually the spectra require a long accumulation. For most biomolecules, i.e. proteins and nucleic acids, the difference in absorbance (scattering) between the two configurations is a few orders of magnitude smaller than the corresponding unpolarized absorbance (scattering).

Most of well known advantages of matrix isolation spectroscopy are related to studies of short-lived molecules and free radicals, but also should be listed: the elimination of hot transitions from the thermally populated vibrational states, sharpness and significantly reduced overlap of the spectral bands, the possibility to study thermally unstable species. In the gas phase, the molecules can be observed unhindered by the environment, and very high-resolution spectra can be used to determine all of the physical properties of a molecular system [1].

1.2 Biological Application

Both IR and Raman spectroscopy can provide important information about the composition and complexity of biological samples. First biological applications of IR spectroscopy were limited to dried plant material and related products because of strong dipole moment of water. The development of Fourier transform (FT) methods pushed the usage of both spectroscopic techniques dramatically. Also better signal-to-noise ratio and shorter time of spectrum acquisition resulted in new applications of these methods. In particular, ATR techniques have improved rapid IR measurements of most liquids such as edible oils, essential oils and solvent extracts of various plant tissues. High accuracy of IR measurements gives an alternative to the conventional methods such as titration and, as a result, IR is most commonly applied for the analysis of dried or non-aqueous plant materials. On the other hand, water has weak Raman scattering properties and consequently Raman methods are more suitable for *in situ* studies of fresh plant materials. Contrary to IR, Raman spectroscopy does not require special sample preparation and optical transparency. Samples can be analyzed directly wet or dry and in many cases nondestructively. Although Raman scattering is weaker than IR absorption, in some cases samples containing lower analyte concentration can be investigated. In particular, two techniques, i.e. resonance Raman and SERS, result in a significant sensitivity enhancement.

Since each functional group in a molecule contributes more or less to the spectral output, vibrational spectra of plant material are usually very complex. The accurate results from spectroscopic analysis can be difficult because of the fact that overlapping and mixing of various vibrational modes occur. Consequently, for many years IR spectroscopy in the middle spectral range (MIR) was used in agricultural studies

only as a qualitative technique for identification and verification of unknown pure substances isolated from extracts or distillates [4]. Nowadays, due to better equipment and modern techniques, FT-IR spectroscopy has become a powerful tool for elucidating the structure, physical properties, and interaction of various plant components. Raman spectra support IR analysis however due to weak scattering some compounds can be difficult to identify.

Vibrational spectroscopy methods can be applied for investigation of primary and secondary plant metabolites. Primary metabolites are defined as those plant constituents, which are essential for the life of plants. It means that they are directly involved in plants normal growth, development, and reproduction. The most important representatives of this group are proteins, lipids, and carbohydrates.

Contrary to primary metabolites, secondary plant metabolites are usually unique to individual plant species and occur in low concentrations. Furthermore, they are not essential to plant cell survival and their importance is related to ecological aspects such as defense against predators, parasites and diseases, interspecies competition, and reproductive processes (colors, smells, etc.). Among a great number of secondary metabolites, the most common are phenolic compounds, terpenoids, alkaloids, polyacetylenes, nitrile compounds, iridoids and chlorophylls. Although their very low concentration in plant material, resonance Raman and FT-Raman spectroscopies provide good spectra with no or little fluorescence [5, 6]. For instance, fruit ripening of various species can be followed by Raman spectroscopy due to the fact that some carotenoids decline when other are accumulated during the ripening process.

Apart from identification of primary and secondary plant metabolites, vibrational spectroscopy gives an opportunity to follow distribution of the metabolites and other plant components simultaneously and directly in plant material (*in situ*). Such distribution studies can be performed using Raman mapping and IR imaging. By combining these techniques with microscopy, molecular information can be obtained with high spatial resolution. There are numerous examples of the usage of *in situ* Raman mapping technique for investigation of plant metabolites distribution in cells, tissues, and whole parts of plants. One of the first Raman mapping studies in plant research was performed on flax stem tissue gaining information on major components (cellulose, lignin, polysaccharides) in different tissue types [7]. The potential of Raman mapping combined with confocal microscope has been shown for the components of cell walls of wood with a high spatial resolution (below 1 μm) [8]. Furthermore, application of polarized laser light provided an insight into changes of orientation of the cell wall constituents [9].

High spatial resolution of Raman microscopy enables characterization of individual compounds in single plant cells. For instance, lipid droplets in xylem [10] and spherical storage compounds in parenchyma cells [8] have been detected in woody plants. Raman spectroscopy can be also applied for distribution studies in single plant cells when imaging techniques are used. Moreover, a single carotenoid crystal has been detected directly in a carrot cell [11]. The cellular distribution of carotenoids with the focus on the individual compounds (β-carotene or astaxanthin) in single cells of various algae was performed [12, 13].

Fast development of FT-IR techniques enabled extensive application of IR imaging for distribution studies of main components in the plant material. Several works have been focused on changes in wood tissues including cell wall modification during maturation [14], fungi degradation [15], steam treatment [16], and transgenic modification [17] to show tissue specific accumulation/degradation patterns. IR imaging has been applied for the analysis of wheat endosperm cell wall composition changes under the impact of environmental conditions [18]. Differentiation between wheat cultivars based on variety of endosperm cell walls [19] was demonstrated. IR imaging technique can be more successful than conventional chemical methods in indicating differences between similar chemotypes, e.g. between wild and mutant type of Arabidopsis [20].

High spatial resolution of IR imaging can be achieved by using a synchrotron source. Synchrotron radiation-based (SR) FT-IR imaging has already been successfully used to examine plant tissue in, e.g. cereals [21]. However, application of FPA detector in IR microspectroscopy has been shown to achieve similar results to synchrotron measurements at the single cell level [22]. Finally, IR imaging based on FPA detector can be applied for bigger areas of a plant such as whole root sections and leaves to observe tissue specific distribution of the studied compounds with comparable spatial resolution and signal-to-noise ratio.

Results from Raman mapping or IR imaging are usually analyzed by comparing integral intensity or by chemometric methods such as Principal Component Analysis (PCA) and Cluster Analysis (CA) to obtain two-dimensional maps and images.

Apart from qualitative analysis of plant material, vibrational spectroscopy is a powerful tool for quantitative studies of individual plant components. This method does not require extraction or separation of the constituents in contrary to widely used chromatographic methods. Vibrational methods allow the quantification of new samples directly from spectral data. Thus, IR and Raman spectroscopy can be used as fast and nondestructive calibration methods for quantification of relevant component contents in plants and related products. Linear calibration based on vibrational spectra and chromatographic data can be used for reliable prediction of the product content as well as for the efficient selection of high-quality single products in industry. It seems to be a powerful tool for a rapid and low-cost alternative quality control method of food production processes.

1.3 Biomedical Application

In the biomedical field it is very desirable to develop innovative and widely used techniques to study the uptake and distribution of bioactive substances. The optical methods do not require any additional labeling or special photophysical properties of the investigated sample. The conventional imaging techniques are often limited by the insufficient sensitivity, specificity and spatial resolution, however the vibrational spectroscopy can be applied to study even single animal cells and various tissues.

Vibrational spectroscopy combined with microscopy has become a powerful diagnostic tool in the biomedical applications. Raman and infrared absorption

spectroscopy appear to be an innovative, powerful, sensitive and non-invasive methods to study the processes taking place inside the cells. IR spectroscopy monitors the level of the main components e.g. proteins, lipids, DNA/RNA, whereas Raman spectroscopy is useful for the assessment of the overall molecular constitution of biological samples, including proteins, nucleic acids, lipids, carbohydrates, heme, carotenoids and inorganic crystals.

The beginnings of biomedical applications of vibrational spectroscopy goes back to the first half and turn of the twentieth century [23–27]. In 1949 Elkan Blout et al. have published one of the first paper where the first spectroscopic experiments on microtomed tissue sections using infrared spectroscopy to study tissue were presented [23]. A normal tissues, neoplastic tissues and important constituents of tissues have been empirically investigated using IR spectroscopy also over 60 years ago by Donald L. Woenley [25]. Recent technological developments have spurred increased interest in Raman microspectroscopy as a sophisticated analytical tool for biomedical applications. The literature repeatedly emphasized the potential impact and importance of these techniques in the application of modern scientific research in the field of tissue engineering and research on new drugs.

The ability to detect diseases or dysfunctions rapidly, non-invasively and unequivocally has multiple benefits. Fast and powerful methods of diagnosis in the initial state of disease allow for early intervention of therapeutic strategies and significant reduction in mortality and morbidity. Detection, identification and tracking of characteristic biochemical markers of disease could be used to monitor the progression of therapy.

Current biomedical studies with application of vibrational spectroscopy are focused on biomedical samples obtained from patients or laboratory animals during surgery, biopsy or *postmortem*, and on cultured cells as well. In recent years there has been a significant increasing interest in the *in vitro* analysis of cells and extracellular matrix components in tissues and *ex vivo* analysis of animal and human tissues by spectroscopic methods for diagnostic purposes. Vibrational spectroscopy opens completely new possibilities for monitoring the content of many chemicals in cells or tissues at the same time at high level of selectivity and resolution. The desire of scientists to have an insight into the biochemical composition of a single cell in an easier, faster and, most importantly, without the labeling in contrast to staining methods, has initiated a number of studies with the use of spectroscopic techniques.

Raman microspectroscopy is able to detect small biochemical changes and their distributions at sub-cellular level [28]. The possibilities to study samples in the aqueous environment allows the investigation of living cells and tracking of changes in their interior under the action of various factors, e.g. monitoring the uptake of drugs, nanoparticles and bioactive compunds, as well as non-chemical factors [29–32].

Many disease syndromes begin at subcellular level, and actually only a few currently used non-invasive techniques allow the study of selected sub-cellular structures in a selective, sensitive and free-labels way. Raman measurements with a spatial resolution ca. ~300 nm allow detection of such small structures as nucleolus, nucleoli, mitochondria, lipid droplets or introduced nanoparticles [33–36]. With the use of confocal Raman spectrometer it is possible to perform three-dimensional imaging

including the position, size and shape of the selected organelles [37]. Changes occurring during the cell cycle, cell death, drug–cell interactions, proliferation differentiation [38–40], or the interaction of cells with various chemicals and materials, can be measured at the biochemical level with high spatial and time resolution.

In the recent years biomedical vibrational spectroscopy has been used to investigate lifestyle diseases, called also diseases of civilization. Lifestyle diseases can include Alzheimer's disease, atherosclerosis, asthma, cancer, chronic liver disease or cirrhosis, type 2 diabetes, heart diseases, metabolic syndrome and many others. Appropriate quantitative approach using Raman or/and infrared spectroscopy combined with chemometric analysis allows to estimate degree of the disease progression.

One direction of the research is focused on the origin, development, treatment, and prevention of cancer—both at the tissue and cellular level. Cancer is one of the most deadly diseases, which modern medicine does not fully understand and is not able to heal. Early detection of cancer through screening based on imaging is probably the major contributor of reduction in mortality for certain cancers. Future developments using Raman spectroscopy and nanoparticles targeted to tumor biomarkers are promising [41]. Raman spectroscopy combined with confocal microscopy can be used for early detection of cancer [42–45] or the analysis of the tumor morphology [46, 47]. Raman and infrared spectroscopy seems to be a promising tool in the diagnosis of bladder [48, 49], prostate [50–52], stomach [53], larynx [54], tonsil [55], lung [56, 57], breast [58–62] and esophagus [63] cancers, as well as basal cell carcinoma [64–65].

Since the first measurements of living cells were achieved by Puppels et al. [66], Raman microspectroscopy was used as a non-invasive and non-destructive tool for probing single living cancer cell while preserving cell integrity and functions, such as adhesion and proliferating capacities. Application of Raman spectroscopy to cell-drug interaction allows for better and faster exploring of disease mechanisms, dependencies and seeking a cure for cancer.

Alzheimer's disease (AD) is a terminal form of dementia resulting from progressive degeneration of the neurons. So far its causes are unknown. Some papers suggest that application of FT-IR spectroscopy could be used in the pathological diagnosis of AD for the classifcation of AD brain tissue (e.g. grey matter), neurotic plaques in the brains, structural characterization β-sheet structure in amyloid fibrils, tau protein and AD-paired helical fragments [67]. Nevertheless because of the penetration depth of infrared light in the tissue (approx. 10 μm) infrared absorption spectroscopy cannot be used for non-invasive diagnosis of AD *in vivo*. For the diagnosis and monitoring of Alzheimer's disease near-IR Raman spectroscopy is a potentially attractive technique because in this system light can penetrate quite deeply into tissue. Mizuno et al. [68], reported a non-destructive examination of human brain tissues and several kinds of brain tumors using near-infrared excited Fourier transform (NIR-FT) Raman spectroscopy. NIR-FT Raman spectra of AD brain tissue show distinct differences from normal tissue spectra that can be used to distinguish AD from normal brain.

A leading cause of death affecting almost one third of humans in developed countries is atherosclerosis. Atherosclerosis is a chronic disease involving degenerative-

productive changes in the intima and central arteries, mainly in the aorta, coronary arteries and cerebral arteries, rarely in the extremities, which leads to a reduction in the light of the arteries and reduce their elasticity by the local deposits of cholesterol, its esters and other lipids [69]. At present, most of the diagnostic techniques commonly used to study patients with atherosclerotic vascular disease (e.g. angiography, nuclear magnetic resonance imaging, electron beam computed tomography or intravascular ultrasound) do not assess the biochemical composition of the vessel wall. Some studies have revealed that Raman and infrared spectroscopy, compared to histopathological analysis as a gold standard, is capable of identifying and classifying the different types of tissues found in the atherosclerotic process of artery *postmortem* [70]. The main purpose of studies with vibrational spectroscopic technique is ability of using label-free optical microscopy to characterize, and thus enable quantitative analyses of different atherosclerotic lesion types. Imaging and quantitative analysis of atherosclerotic lesions was possible by using Coherent Anti-Stokes Raman Scattering (CARS) nonlinear optical microscopy [71]. One of the first real-time investigation utilizing Raman spectroscopy to examine human atherosclerosis *in vivo*, (during femoral bypass procedures) and also to demonstrate the sensitivity of this technique to identify spectroscopic features associated with plaque vulnerability was presented by Motz et al. [72]. Because in vivo Raman spectroscopy does not require the removal of tissue, its success might open several new avenues of research.

Diabetes mellitus (DM) is an increasingly common metabolic disorder triggered by the absence or deficiency of insulin, insulin resistance or by defect in insulin secretion and/or insulin action. As a result hyperglycaemia and profound perturbations in carbohydrate, fat and protein metabolism are observed. These biochemical alterations of cellular metabolism in diabetes include changes in biochemical profile of vascular wall as has been already detected by FT-IR spectroscopy. Variations in spectral parameters can be used in early diagnosis of DM, which may also stimulate the development of patient monitoring devices in future, what can be important in early detection of metabolic disorders. There are only a few reports of vibrational spectroscopy focused on diabetic samples [73] and measurements of glucose concentrations in blood [74]. FT-IR spectroscopy is very informative to differentiate diabetic tissues from healthy ones at the molecular level [75].

Raman spectroscopy is a valuable tool to investigate and follow an oxidative stress, which is one of the potential marker of lifestyle diseases. Oxidative stress is associated with increased production of reactive oxygen species (ROS) and impaired antioxidant mechanisms. Direct observation of reactive oxygen species (ROS) and oxidative stress using vibrational spectroscopy is not possible, but information about ROS can be obtained through the analysis of the effects of ROS on cells and tissues. Krafft et al. [36] presented a novel description of stress-induced changes at subcellular level (nucleus, cytoplasm, vesicles, inclusion bodies, and the peripheral membranes) and apoptosis of cells based on Raman microspectroscopy imaging. Raman spectroscopy has been shown to be also useful in determining correlation between carotenoid antioxidants in living human tissues and risk for malignancies or other diseases associated with oxidative stress, such as e.g. cancer,

atherosclerosis or Alzheimer's disease [76, 77]. Development of the methodology should allow for the diagnostic assessment of biochemical changes caused by the oxidative stress.

References

1. Chalmers JM, Griffiths PR (eds) (2002) Handbook of vibrational spectroscopy, vol 1–5. Wiley, Chichester
2. Guerrini L, Jurasekova Z, Domingo C, Pérez-Méndez M, Leyton P, Campos-Vallette M, Garcia-Ramos JV, Sanchez-Cortes S (2007) Importance of metal–adsorbate interactions for the surface-enhanced Raman scattering of molecules adsorbed on plasmonic nanoparticles. Plasmonics 2:147–156
3. Lombardi JR, Birke RL (2009) A unified view of surface-enhanced Raman scattering. Acc Chem Res 42(6):734
4. Colthup NB, Daly LH, Wiberley SE (1990) Introduction to infrared and Raman spectroscopy. Academic Press, San Diego
5. Ozaki Y, Cho R, Ikegawa K, Muraishi S, Kawauchi K (1992) Appl. Spectrosc 46:1503
6. Withnall R, Chowdhry BZ, Silver J, Edwards HGM, de Oliveira LFC (2003) Spectrochim Acta A 59:2207
7. Himmelsbach DS, Khahili S, Akin DE (1999) Vib Spectrosc 19:361–367
8. Gierlinger N, Schwanninger M (2006) Chemical imaging of poplar wood cell walls by confocal Raman microscopy. Plant Physiol 140(4):1246–1254
9. Gierlinger N, Burgert I (2006) Secondary cell wall polymers studied by Confocal Raman microscopy: spatial distribution, orientation and molecular deformation. New Zeal J For Sci 36(1):60–71
10. Rösch P, Schneider H, Zimmermann U, Kiefer W, Popp J (2004) *In situ* Raman investigation of single lipid droplets in the water-conducting xylem of four woody plant species. Biopolymers 74(1–2):151–156
11. Baranska M, Baranski R, Grzebelus E, Roman M (2011) Vib Spectrosc 56:166–169
12. Kaczor A, Turnaub K, Baranska M (2011) Analyst 136:1109–1112
13. Collins AM, Jones HDT, Han D, Hu Q, Beechem TE, Timlin JA (2011) PLoS ONE 6(9):e24302
14. Gou J-Y, Park S, Yu X-H, Miller LM, Liu C-J (2008) Planta 229:15–24
15. Fackler K, Stevanic JS, Ters T, Hinterstoisser B, Schwanninger M, Salmén L (2010) Enzyme Microb Tech 47:257–267
16. Yin Y, Berglund L, Salmen L (2011) Biomacromolecules 12:194–202
17. Labbe N, Rials TG, Kelley SS, Cheng Z-M, Kim J-Y, Li Y (2005) Wood Sci Technol 39:61–77
18. Toole GA, Wilson RH, Parker ML, Wellner NK, Wheeler TR, Shewry PR, Mills ENC (2007) Planta 225:1393–1403
19. Barron C, Parker ML, Mills ENC, Rouau X, Wilson RH (2005) FTIR imaging of wheat endosperm cell walls in situ reveals compositional and architectural heterogeneity related to grain hardness. Planta 220:667–677
20. Gorzsas A, Stenlund S, Persson P, Trygg J, Sundberg B (2011) Plant J 66:903–914
21. Wetzel DL, Eilert AJ, Pietrzak LN, Miller SS, Sweat JA (1998) Ultraspatially-resolved synchrotron infrared microspectroscopy of plant tissue in situ. Cell Mol Biol 44:145–167
22. Heraud P, Caine S, Sanson G, Gleadow R, Wood BR, McNaughton D (2007) New Phytol 173:216–225
23. Blout Elkan R, Mellors Robert C (1949) Infrared spectra of tissues. Science 110:137–138
24. Schwarz H (1951) Infrared spectra of tissues. Appl Spectrosc 6:15–18

25. Woernley D (1952) Infrared absorption curves for normal and neoplastic tissues and related biological substances. Cancer Res 12:516–523
26. Mellors R (1953) Microscopy. I: a review. Cancer Res 13:101–118
27. Bordner RH, Kabler PW, Kenner B et al (1956) Bacterial identification by infrared spectrophotometry. J Bacteriol 72:593–603
28. Short KW, Carpenter S, Freyer JP et al (2005) Raman spectroscopy detects biochemical changes due to proliferation in mammalian cell cultures. Biophys J 88:4274–4288
29. Miljkovic M, Romeo M, Matthäus C et al (2009) Infrared Microspectroscopy of Individual Human Cervical Cancer (HeLa) Cells Suspended in Growth Medium. Biopolymers 74:172–175
30. Wood BR, Tait B, McNaughton D (2001) Micro-Raman characterisation of the R to T state transition of haemoglobin within a single living erythrocyte. Biochim Biophys Acta 1539:58–70
31. Kuimova MK, Chan KLA, Kazarian SG (2009) Chemical imaging of live cancer cells in the natural aqueous environment. Appl Spectrosc 63:164–171
32. Bonnier F, Byrne HJ (2012) Understanding the molecular information contained in principal component analysis of vibrational spectra of biological systems. Analyst 137:322–332
33. Tang H, Yao H, Wang G et al (2007) NIR Raman spectroscopic investigation of single mitochondria trapped by optical tweezers. Opt Express 15:12708–12716
34. Lasch P, Pacifico A, Diem M (2002) Spatially resolved IR microspectroscopy of single cells. Biopolymers 67:335–338
35. Lasch P, Boese M, Pacifico A et al (2002) FT-IR spectroscopic investigations of single cells on the subcellular level. Vib Spectrosc 28:147–157
36. Krafft C, Knetschke T, Funk RHW et al (2006) Studies on stress-induced changes at the subcellular level by Raman microspectroscopic mapping. Anal Chem 78:4424–4429
37. Majzner K, Kaczor A, Kachamakova-Trojanowska N et al (2013) 3D confocal Raman imaging of endothelial cells and vascular wall: perspectives in analytical spectroscopy of biomedical research. Analyst 138:603–610
38. Holman HY, Martin MC, Blakely EA et al (2000) IR spectroscopic characteristics of cell cycle and cell death probed by synchrotron radiation based Fourier transform IR spectromicroscopy. Biopolymers 57:329–335
39. Zelig U, Kapelushnik J, Moreh R et al (2009) Diagnosis of cell death by means of infrared spectroscopy. Biophys J 97:2107–2114
40. Liu KZ, Jia L, Kelsey SM et al (2001) Quantitative determination of apoptosis on leukemia cells by infrared spectroscopy. Apoptosis 6:269–278
41. Fass L (2008) Imaging and cancer: a review. Mol Oncol 2:115–152
42. Harris AT, Rennie A, Waqar-Uddin H et al (2010) Raman spectroscopy in head and neck cancer. Head Neck Oncol 2:26
43. Kendall C, Isabelle M, Bazant-Hegemark F et al (2009) Vibrational spectroscopy: a clinical tool for cancer diagnostics. Analyst 134:1029–1045
44. Brozek-Pluska B, Musial J, Kordek R et al (2012) Raman spectroscopy and imaging: applications in human breast cancer diagnosis. Analyst 137:3773–3780
45. Abramczyk H, Placek I, Bro B (2008) Human breast tissue cancer diagnosis by Raman spectroscopy. Spectroscopy 22:113–121
46. Bergner N, Krafft C, Geiger KD et al (2012) Unsupervised unmixing of Raman microspectroscopic images for morphochemical analysis of non-dried brain tumor specimens. Anal BioAnal Chem 403:719–725
47. Kast R, Serhatkulu G, Cao A (2008) Raman spectroscopy can differentiate malignant tumors from normal breast tissue and detect early neoplastic changes in a mouse model. Biopolymers 89:235–241
48. Draga ROP, Grimbergen MCM, Vijverberg PLM et al (2010) In vivo bladder cancer diagnosis by high-volume Raman spectroscopy. Anal Chem 82:5993–5999

49. Barman I, Dingari NC, Singh GP et al (2012) Selective sampling using confocal Raman spectroscopy provides enhanced specificity for urinary bladder cancer diagnosis. Anal Bioanal Chem 404:3091–3099
50. Harvey TJ, Gazi E, Henderson A et al (2009) Factors influencing the discrimination and classification of prostate cancer cell lines by FTIR microspectroscopy. Analyst 134:1083–1091
51. Bhargava R (2007) Towards a practical Fourier transform infrared chemical imaging protocol for cancer histopathology. Anal Bioanal Chem 389:1155–1169
52. Bhargava R, Fernandez DC, Hewitt SM et al (2006) High throughput assessment of cells and tissues: Bayesian classification of spectral metrics from infrared vibrational spectroscopic imaging data. Biochim Biophys Acta 1758:830–845
53. Teh SK, Zheng W, Ho KY et al (2008) Diagnostic potential of near-infrared Raman spectroscopy in the stomach: differentiating dysplasia from normal tissue. Br J Cancer 98:457–465
54. Stone N, Stavroulaki P, Kendall C et al (2000) Raman spectroscopy for early detection of laryngeal malignancy: preliminary results. Laryngoscope 110:1756–1763
55. Lyng FM, Faoláin EO, Conroy J et al (2007) Vibrational spectroscopy for cervical cancer pathology, from biochemical analysis to diagnostic tool. Exp Mol Pathol 82:121–129
56. Huang Z, McWilliams A, Lui H et al (2003) Near-infrared Raman spectroscopy for optical diagnosis of lung cancer. Int J Cancer 107:1047–1052
57. Kaminaka S, Yamazaki H, Ito T et al (2001) Near-infrared Raman spectroscopy of human lung tissues: possibility of molecular-level cancer diagnosis. J Raman Spectrosc 32:139–141
58. Eckel R, Huo H, Guan H-W et al (2001) Characteristic infrared spectroscopic patterns in the protein bands of human breast cancer tissue. Vib Spectrosc 27:165–173
59. Walsh MJ, Kajdacsy-Balla A, Holton SE et al (2012) Attenuated total reflectance Fourier-transform infrared spectroscopic imaging for breast histopathology. Vib Spectrosc 60:23–28
60. Bitar Carter RA, Martin AA, Netto MM et al (2004) FT-Raman spectroscopy study of human breast tissue. In: Mahadevan-Jansen A, Sowa MG, Puppels GJ et al (eds) Proc. SPIE 5321, Biomedical vibrational spectroscopy and biohazard detection technologies. Bellingham, WA, pp 190–197
61. Abramczyk H, Brozek-Pluska B, Surmacki J et al (2011) The label-free Raman imaging of human breast cancer. J Mol Liq 164:123–131
62. Abramczyk H, Brozek-Pluska B, Surmacki J et al (2012) Raman "optical biopsy" of human breast cancer. Prog Biophys Mol Biol 108:74–81
63. Shetty G, Kendall C, Shepherd N et al (2006) Raman spectroscopy: elucidation of biochemical changes in carcinogenesis of oesophagus. Br J Cancer 94:1460–1464
64. Larraona-Puy M, Ghita A, Zoladek A et al (2012) Development of Raman microspectroscopy for automated detection and imaging of basal cell carcinoma. J Biomed Opt 14:054031
65. Nijssen A, Bakker Schut TC, Heule F et al (2002) Discriminating basal cell carcinoma from its surrounding tissue by Raman spectroscopy. J Invest Dermatol 119:64–69
66. Puppels GJ, De Mul FF, Otto C et al (1990) Studying single living cells and chromosomes by confocal Raman microspectroscopy. Nature 347:301–303
67. Fabian H, Choo LI, Szendrei GI et al (1993) Infrared spectroscopic characterization of Alzheimer plaques. Appl Spectrosc 47:1513–1518
68. Mizuno A, Kitajima H, Kawauchi K et al (1994) Near-infrared Fourier transform Raman spectroscopic study of human brain tissues and tumours. J Raman Spectrosc 25:25–29
69. Tabas I (2008) Lipids and atherosclerosis. In: Vance J, Vance D (eds) Biochemistry of lipids, lipoproteins and membranes, 5th ed. Elsevier, Oxford, pp 579–604
70. Peres MB, Silveira L, Zângaro RA et al (2011) Classification model based on Raman spectra of selected morphological and biochemical tissue constituents for identification of atherosclerosis in human coronary arteries. Lasers Med Sci 26:645–655
71. Wang H-W, Langohr IM, Sturek M et al (2009) Imaging and quantitative analysis of atherosclerotic lesions by CARS-based multimodal nonlinear optical microscopy. Arterioscler Thromb Vasc Biol 29:1342–1348
72. Motz JT, Fitzmaurice M, Haka AS et al (2006) In vivo Raman spectral pathology of human atherosclerosis and vulnerable plaque. J Biomed Opt 11:1–9

73. Toyran N, Lasch P, Naumann D et al (2006) Early alterations in myocardia and vessels of the diabetic rat heart: an FTIR microspectroscopic study. Biochem J 397:427–436
74. Khalil OS (1999) Spectroscopic and clinical aspects of noninvasive glucose measurements. Clin Chem 45:165–177
75. Petibois C, Gionnet K, Gonçalves M et al (2006) Analytical performances of FT-IR spectrometry and imaging for concentration measurements within biological fluids, cells, and tissues. Analyst 131:640–647
76. Ermakov IV, Ermakova MR, McClane RW et al (2001) Resonance Raman detection of carotenoid antioxidants in living human tissues. Opt Lett 26:1179–1181
77. Puppels GJ, Garritsen HS, Kummer JA et al (1993) Carotenoids located in human lymphocyte subpopulations and natural killer cells by Raman microspectroscopy. Cytometry 14:251–256

Chapter 2
Integrating Optical Spectroscopy and Chemometric Methods

Karolina Jagiełło, Anita Sosnowska, Jan Mazerski and Tomasz Puzyn

Abstract In this Chapter, we describe the usage of several chemometric and numerical techniques to analyse of UV-Vis sets of spectra. The fundamentals of each technique are briefly presented with examples of its applications. This approach allows obtaining deeper insight in studied system. These methods can be used not only to analyze aggregation process, as it was presented in the Chapter, but also to study the interaction between small ligands and macromolecules, such as DNA. Determination of the number of formed complexes and the binding constant of interaction ligand/macromolecules can be received with this methodology.

Keywords Chememetrics • UV-Vis spectra • Aggregation • Imidazoacridines

2.1 Introduction

In chemistry, we have been often dealing with a set of measured data, which are in fact a mixture of the information and the noise. In many cases, the magnitude of the noise is as much great as the information that significantly hinders the ability to find the interesting results in multidimensional data sets [1]. The solutions of such problems are proposed by chemometrics. By means of chemometric techniques we are able to: (i) delete as much noise as possible from the data sets; and, (ii) extract as much information as possible from the multidimensional data [1].

However, to obtain reliable results of chemometrics analysis some special rules should be applied.

The chemical/analytical problem that one would like to solve, has to be defined precisely before starting the experiment. Problems solved by means of chemometric methods can be grouped into four main families: making the

T. Puzyn (✉) · K. Jagiełło · A. Sosnowska
Laboratory of Environmental Chemometrics, Faculty of Chemistry, University of Gdansk, Gdansk, Poland
e-mail: puzi@qsar.eu.org

J. Mazerski
Department of Pharmaceuticals Technology and Biochemistry, Chemical Faculty, Gdansk University of Technology, Gdansk, Poland

M. Baranska (ed.), *Optical Spectroscopy and Computational Methods in Biology and Medicine,* Challenges and Advances in Computational Chemistry and Physics 14,
DOI 10.1007/978-94-007-7832-0_2, © Springer Science+Business Media Dordrecht 2014

visualisation of multivariate data set; searching the relationships between sets of data; recognition of internal data structure and making the classification [2]. Each of them needs adequate type of data sets and usage the appropriate chemometric approaches.

The next step is to design carefully the experiment and then perform measurements. The choice of the experimental design depends on the type of information required to verify the research hypothesis, which in fact is the solution of the defined problem. It is indicated to perform as little trials as possible, but selected samples should be representative [1, 2]. In such way, we do not only save time and cost of the whole project but also let make to increase significantly the quality of the results.

Collected data should undergo multi-aspect control procedures to: (i) identify the outliers; (ii) determine of the variable distribution that in some cases indicate the necessity to perform data transformation; and, (iii) determine of interdependences between variables. The main goal of this procedure is to eliminate errors and shorten the time necessary to get reliable results [2].

Only data prepared according to described above procedures can be analysed by means of chemometric techniques. The main approaches of chemometrics data analysis are presented in Fig. 2.1.

Summing up, **chemometrics is the field of science dealing with extracting valuable information from multidimensional data sets by means of mathematical and statistical methods**. However, appropriate usage of chemometrics requires thinking about its application in each step of experiment: from the problem defining to the analysis of obtained results.

2.2 Chemometrics and Numerical Approach to a Set of Spectra

One of a typical problem in absorption spectroscopy is to determine an influence of experimental conditions, such as temperature, ionic strength of the medium or pH, on the electronic spectrum of tested compounds. The experimental condition may affect intensity as well as shape of the spectrum. To detect such changes a set of spectra registered at different experimental condition should be analysed.

From chemometrics point of view each spectrum could be treated as a one-dimensional matrix. Absorptions at different wavelengths are elements of this matrix. A set of spectra registered at different condition forms two-dimensional matrix **X** [3, 4].

2.2.1 Preprocessing of Spectra

Registered sets of spectra, transformed to molar extinction scale according to the Lambert-Beer Eq. (2.1) have to be preprocessed.

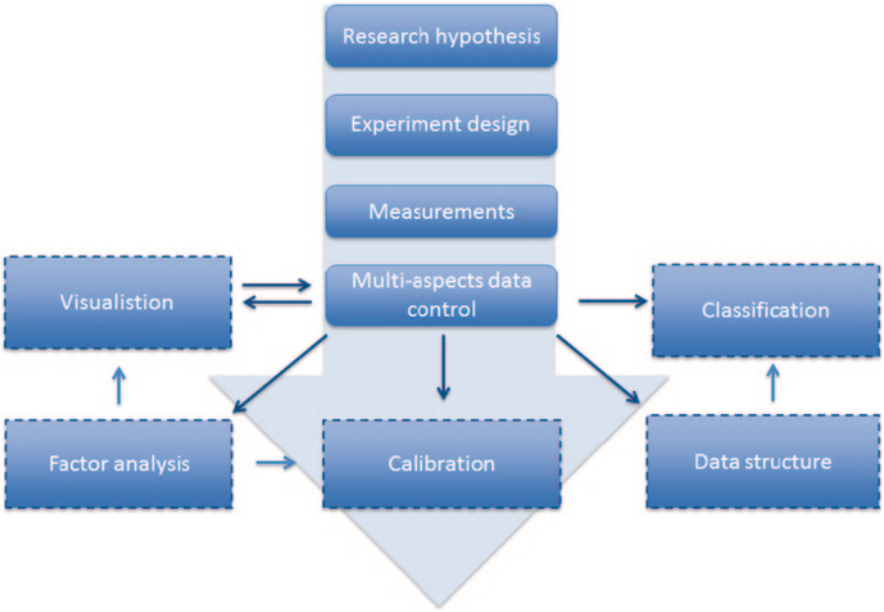


Fig. 2.1 The overview of chemometrics methods

$$A_\lambda = \varepsilon_\lambda cl \tag{2.1}$$

A_λ absorbance at a given wavelength, expressed in nanometers,
ε_λ the molar extinction coefficient at a given wavelength,
c molar concentration of compound,
l optical path length expressed in centimeters

The first steps of these operations include: baseline correction and dilution correction. Several techniques applied to perform baseline correction are summarized by Kohler et al. [5].

Obtain, corrected and transformed results are organize into matrix **X**, presented in Table 2.1 consists n rows (wavelengths, λ) and m columns (spectra obtained for the m increasing concentrations of the compound or spectra registered at different temperatures, pH, etc.). Each element of this matrix corresponds to extinction values calculated for each sample in following wavelengths.

For each column of matrix, the mean values and standard deviation values can be calculated for whole range of wavelengths or for selected range of wavelength, for example range characteristic for research chromophore. These two parameters can be used for measuring changes in intensity of the spectra [3, 4]. To minimalize influence of changes in intensity of spectra, and in this way study only changes into its shapes, sets of spectra should be transformed. The most popular transformations are centring and autoscaling.

Table 2.1 Matrix X obtain as a result of 1st step preprocessing operations of set of UV-Vis spectra

Wavelength	Sample 1	Sample 2	…	Sample m
λ_1	$\varepsilon_{1,1}$	$\varepsilon_{1,2}$	…	$\varepsilon_{1,m}$
λ_2	$\varepsilon_{2,1}$	$\varepsilon_{2,2}$	…	$\varepsilon_{2,m}$
.	.	.	…	.
λ_n	$\varepsilon_{n,1}$	$\varepsilon_{n,2}$	…	$\varepsilon_{n,m}$

Centring the data is the linear transformation performs to make all variables coincide with the beginning of the coordinate system. It is done by subtracting from each value x_{ij} of data organized into matrix **X** corresponding mean value of jth variable to obtain new matrix **X′**, according to formula [1]:

$$x'_{ij} = x_{ij} - \mu_j \tag{2.2}$$

The autoscaling is made by diving by centered data by the standard deviation of the jth variables according to formula (2.2). As a result, we obtained the normalized matrix **Z** [2]:

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j} \tag{2.3}$$

where:

x_{ij} molar extinction coefficient for the i-th wavelength of the j-th spectrum
μ_j the average value of the molar extinction coefficients of the j-th spectrum
σ_j standard deviation of the molar extinction coefficients of the j-th spectrum.

Spectrum after standardization can be treated as a vector of unit length. A set of spectra creates a bunch of vectors which have common origin in the k-dimensional hyperspace, where k is the number of spectral forms present in the analyzed samples. A bunch of vectors for recorded spectra is limited by vectors representing the spectra of pure ingredients. An example for two component mixture (k = 2) is shown in Fig. 2.2.

2.2.2 Internal Order of a Matrix and Number of Variety Sources

The Principal Component Analysis (PCA) [6] is a special example of projection pursuit techniques, in which variance is used as a projection index. PCA is mainly used for modeling, compressing and visualizing multidimensional data [7–10]. Application PCA in the relationship analysis involves two basic tasks: graphical

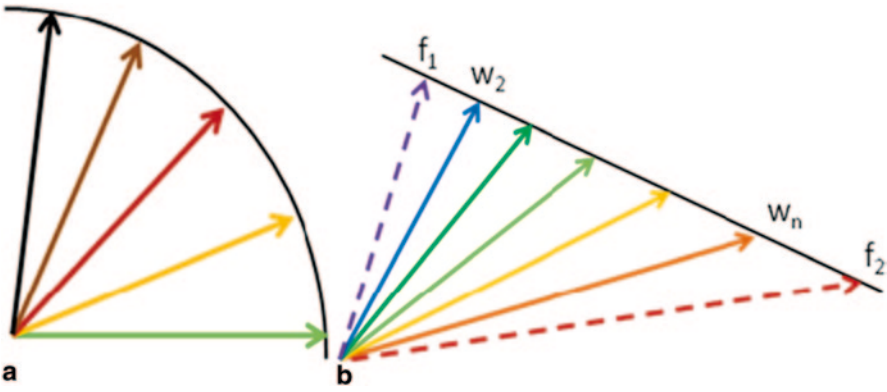


Fig. 2.2 Vectors of the two component mixture in the space of two main components: **a** standardized spectrum. **b** centered spectrum

presentation of the dimensional dependence and reducing the dimensionality of the problem. In this technique, set of correlation coefficient between variables in multivariate space might be transformed to equivalent set of orthogonal factors. The principal components are new orthogonal variables (which are expressed as a linear combination of original variables) and they maximized description of the data variance.

The PCA analysis allows to determine the internal dimension of the matrix of the standardized spectra. This dimension is equal to the number of spectral forms presented in the samples. If the spectra of individual species are not excessively correlated, it is equal to the number of significant principal components of this matrix. Unfortunately, when the spectral forms have very similar spectra, as is usual in the case of aggregation, the analysis of the residual spectra have to be used to determine the number of spectral species in the samples [3, 4].

The principal components and their loadings permit to complete reconstruction of the standardized spectra matrix according to the equation:

$$\mathbf{Z}_{nm} = \mathbf{P}_{nm}\mathbf{L}_{nm} \tag{2.4}$$

where:

$\mathbf{P}_{nm}$ matrix of m principal components,
$\mathbf{L}_{nm}$ matrix of loadings.

If in the spectra reconstruction we use only j first principal components we obtain following relationship:

$$\mathbf{Z}_{nm} = \mathbf{P}_{nj}\mathbf{L}_{jm} + \mathbf{E}^{(j)} \tag{2.5}$$

$\mathbf{E}^{(j)}$ matrix represents the residual spectra of j order. The residual spectra for $j<k$ have the character of the differences spectra, for $j \geq k$ the residual spectrum represents random noise. Thus, further analysis of the residual spectra allows to specify the correct value of k.

2.2.3 Numerical Spectrum Decomposition Technique and Physicochemical Model of a Process

Applying PCA analysis let us know not only the number of spectral forms presented in the samples, but also the first spectra approximation, however, does not allow to designate individual forms, because the principal components are obtained with the assumption of perfect orthogonality.

Therefore, next step in chemometrical analysis of sets of spectra is to obtain the spectra of indywidauals and estimation of its molar fractions. To reach this goal, **Z** matrix should be expressed as a multiplication of the two other matrices: the matrix of components **B**, in which the columns represent the spectrum of pure spectral forms, and matrix **X**, which columns represent different forms of molar fractions [3, 4].

$$\mathbf{Z} = \mathbf{BX} \tag{2.6}$$

Knowing that the ends of the centered vectors of two components mixtures are on a straight line (Fig. 2.2b), the position of the spectra of pure components on this line is unknown. Therefore, in order to solve the Eq. (2.6) there is need to use the iterative procedure called Numerical Spectrum Decomposition [1]. To determine spectra of the complexes, as well as relative amount of each species, an iterative self-consisting procedure—multivariate curve resolution-alternating least squares (MCR-ALS) has to be used [11]. The most outer vectors w1 i wm participate the most within the spectral forms according to the equation.

$$w_1 = x_{11}f_1 + x_{21}f_2 \quad x_{11} \approx 1, x_{21} \approx 0 \tag{2.7}$$

$$w_m = x_{1m}f_1 + x_{2m}f_2 \quad x_{1m} \approx 1, x_{2m} \approx 0 \tag{2.8}$$

In addition, in the two components mixtures, the molar ratios should not be negative, and the sum of the molar fractions of the two forms must be equal to 1:

$$x_{11} + x_{21} = 1 \tag{2.9}$$

$$x_{1m} + x_{2m} = 1 \tag{2.10}$$

Using the Eq. (2.9) and (2.10) molar fraction x_{11} i x_{2m} can be determined:

$$x_{11} = 1 - x_{21} \tag{2.11}$$

$$x_{2m} = 1 - x_{1m} \tag{2.12}$$

After substituting these molar fractions into formulas (2.11) and (2.12) set of two equations were obtained:

$$\begin{cases} w_1 = x_{11} f_1 + (1 - x_{11}) f_2 \\ w_m = (1 - x_{2m}) f_1 + x_{2m} f_2 \end{cases} \tag{2.13}$$

These equations can be solved due to the spectrum of pure form. Assuming further indications:

$$\alpha = \frac{1}{x_{11}} > 1 \tag{2.14}$$

$$\beta = \frac{1}{x_{2m}} > 1 \tag{2.15}$$

the following formulas are obtained:

$$f_1 = \frac{\alpha w_m - (\alpha - 1)\beta w_m}{1 - (\alpha - 1)(\beta - 1)} \tag{2.16}$$

$$f_2 = \frac{\beta w_m - \alpha(\beta - 1) w_1}{1 - (\alpha - 1)(\beta - 1)} \tag{2.17}$$

Coefficients α and β are matched numerically during fitting a physicochemical model of the process studied using the Nelder–Mead simplex method [12, 13]. Obtaining optimal values of α and β is equivalent to determine the spectra of pure species. Determination of the spectra of pure spectral forms allow to calculate the matrix Eq. (2.6) because of the mole fractions. Molar fraction of the individual components in various conditions obtained by this method, could be used to determine the specific parameters of the model used.

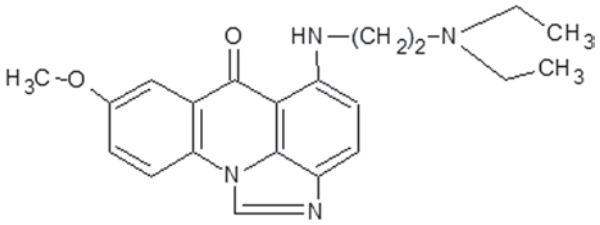


Fig. 2.3 Chemical structure of C-1330

2.3 Influence of Agregation on UV/Vis Spectrum—A Case Study

Acridine belong to a group of polycyclic heteroaromatic compounds and exhibit a broad spectrum of biological activity including antiprotozoal, antibacterial, antiviral and antitumor activity [7, 14]. Imidazoacridinones (IA) derivaties are a group of acridine antitumor drugs synthesized in Department of Pharmaceutical Technology and Biochemistry, Gdańsk University of Technology. The biological activity of the imidazoacridinones has now been extensively investigated. Published data suggested that imidazoacridinone drugs are capable of binding physicochemically to DNA [15, 16]. The nature of these investigation and their relevance to cytotoxic and antitumor properties of IA remains unknown.

Imidazoacrinononones tend to aggregate in diluted solutions, leading to dimer formation, and sometimes even higher order aggregates. Driving forces of this process are hydrophobic interactions. Kinetic information of aggregation process are very helpful to understand molecular interaction such as micelle formation of amphiphilic substances and the binding of small ligands to macromolecules [17–19]. In spite of many studies, the mechanism of compounds self-aggregation seems to be not fully understood because of the experimental conditions, which are different from study to study.

The most useful method to check if some compounds self-aggregate in aqua solution is spectroscopic analysis with increasing drug concentration. The following procedure of results analysis may be used to study the aggregation of various derivatives. An example of such derivative is imidazoacridinon C-1330 (Fig. 2.3), synthesized in Department of Pharmaceutical Technology and Biochemistry, Gdańsk University of Technology [20]. The samples of imidazoacridinon C-1330 were prepare in various concentrations (20 μM to 1 mM) by dilutions. The absorption spectra at different concentration were recorded. Obtained spectra were transformed into the molar scale using Lambert-Beer Eq. (2.1), and then standardized according to Eq. (2.3).

Figure 2.4 represents the set of standardized spectra obtained for C-1330 in different drug concentration. The compound has a strong absorption band in the visible field of the spectrum. Two maximum are observed in the spectra. For the lowest concentration of the compound, they are located at $\lambda = 371$ and $\lambda = 423$ nm. In real spectra, there are two isosbestic points at a wavelength $\lambda = 383$ and $\lambda = 431$ nm.

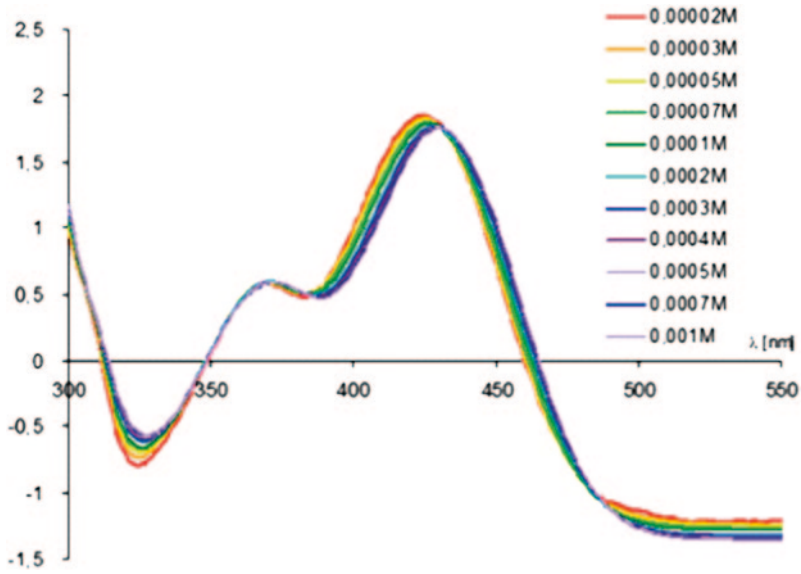


Fig. 2.4 Set of standardized spectra representative for C-1330 in buffer with 5 mM NaCl

Table 2.2 Eigen-values obtained for set of the C-1330 spectra

	No of PC				
	1	2	3	4	5
Eigen-value	9.9827	0.0170	0.0002	0.0001	0.0000
% of variance	99.83	0.16	0.01	0.00	0.00
Cumulative % of variance	99.83	99.99	100.00	100.00	100.00

Numerical analysis of the spectra for C-1330 was started by determining the number of spectral forms present in the solution. Table 2.2 presents the data obtained by using the PCA method. The following values suggest that the first two principal components (explaining 99.99 % of the total variability) are significant.

Similar results were obtained after careful analysis of the residual spectra (Fig. 2.5). The 1st and 2nd order residual spectra are relatively intensive and demonstrate the presence of absorption bands. Experimental noise prevails only in 3rd or higher order residual spectra. This indicates that there are two spectral forms in the solution: monomer and aggregate.

Using the appropriate transformations as it was presented in Chap. 2.3. there were possible to reproduce spectra of pure spectral forms. Figure 2.6 shows such spectra for C-1330. The maximum absorption of the monomer occurs at 423 nm. In the spectra of the aggregate exhibit characteristic batochromic effect with the maximum absorption at the 431 nm. The results of our studies indicate that the presence as well as proportions of particular spectral forms is dependent on the overall concentration of studied compound.

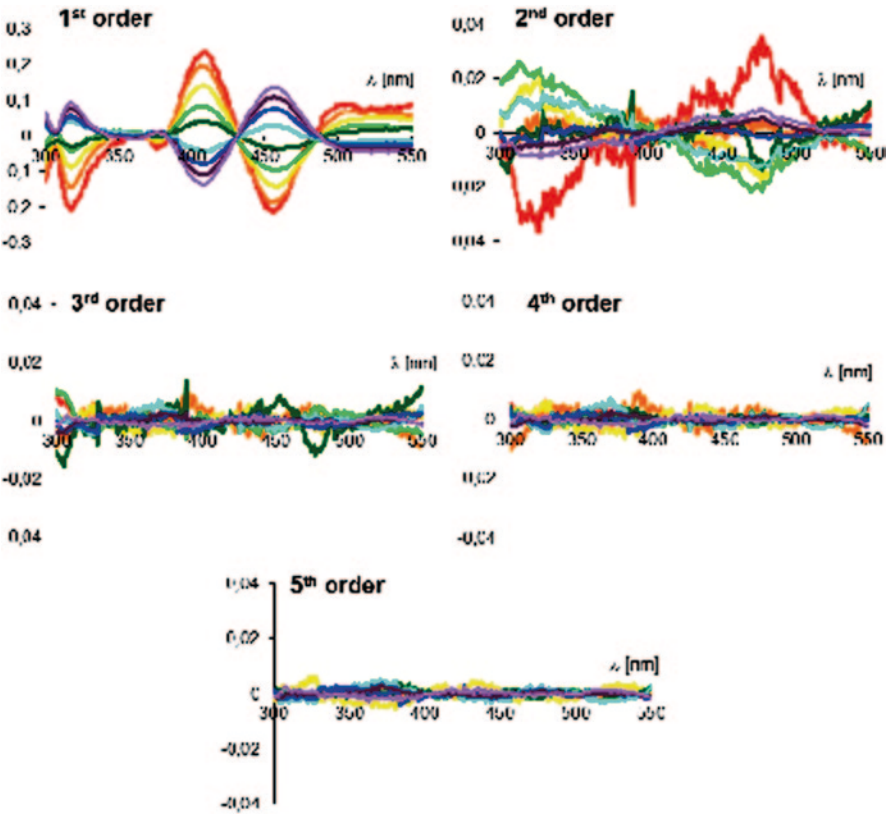


Fig. 2.5 Further residual spectrum obtained for the C-1330 in 5 mM NaCl solution in the PCA method

Obtained molar fraction of the individual components in various concentrations was used to determine the aggregation constant K_A according to the formula:

$$K_{Ai} = \frac{1 - \sqrt{x_{mi}}}{C_{ti} x_{mi}} \tag{2.18}$$

Where

C_{ti} the total concentration of the compound in the i-th solution,
x_{mi} the molar fraction of the monomer calculated from the Eq. (2.6).

The estimation of the m aggregation constant—one for each compound concentrations (for each spectra) were obtained. Than the average value and the standard deviation were calculated: $K_A = (6,327 \pm 0,218) * 10^3 M^{-1}$.

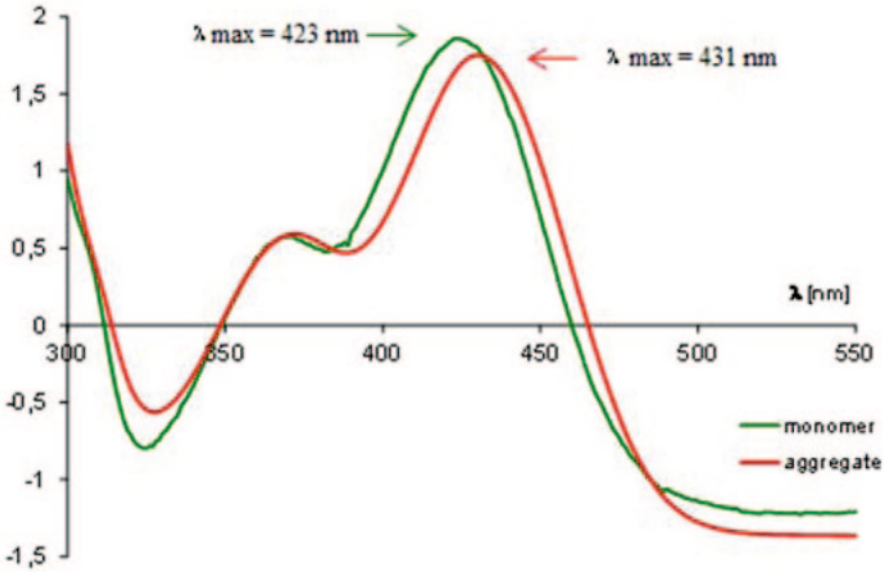


Fig. 2.6 Standardized spectra for monomer (*red line*) and aggregate (*green line*) for C-1330 in 5 mM NaCl

2.4 Conclusions

Application of chemometric and numerical methods can be used to analyse spectroscopic spectra. We presented these applications for the aggregations process that give interesting results even when the spectra of spectral forms presented in the sample differ slightly. The presented approach is universal if the model of the process is known. This methods can be also used to study the interaction between small ligands and macromolecules, such as DNA, for determining the number of formed complexes and the binding constant of interaction ligand/macromolecules. The chemometrics approach to an experimental data allows obtaining deeper insight in studied system. Application of various statistical methods is particular useful when classical simplified assumption is not valid.

Acknowledgements This work was supported by the Polish Ministry of Science and Higher Education, Grant No. DS/8180-4-0171-13

References

1. Kramer R (1998) Chemometric techniques for quantitative analysis. Dekker, New York, United States of America
2. Mazerski J (2006) Introduction to chemometrics. In: Zuba D, Parczewski A (eds) Chemometrics—methods and applications. Institute for Forensic Research Publisher, Krakow, Poland, pp 11–16
3. Slisz M, Mazerski J (2006) Chemometrics analysis of UV-Vis spectra of associated forms of amphotericin B and its derivatives. In: Zuba D, Parczewski A (eds) Chemometrics—methods and applications. Institute for Forensic Research Publisher, Krakow, Poland, pp 307–312
4. Jagiello K, Mazerski J (2006) Spectrophotometric titration of dsDNA ny nitroacridines—chemometrics approach. In: Zuba, D, Parczewski A (eds) Chemometrics—methods and applications. Institute for Forensic Research Publisher, Krakow, Poland, pp 301–306
5. Kohler A, Zimonja M, Segtnan V, Martens H (2009) Standard normal variate, mutliplicative signal correction and extended multiplicative signal correction preprocessing in biospectroscopy.In: Brown SD, Tauler R, Walczak B (eds) Comprehensive chemometrics: chemical and biochemical data analysis, vol 2. Elsevier, The Netherlands, pp 139–163
6. Sharaf MA, Ullman DL, Kowalski BR (1998) Chemometrics. Wiley, New York
7. Wolfe AD (1975) Quinacrine and other acridines. In: Corocoran JW (ed) Antibiotics Vol. III. mechanism of action of antumicrobial and antitumor agents. Springer Berlin, Heidelberg
8. Gemperline P (2006) Practical guide to chemometrics. Taylor, New York
9. Massat DL, Vander Heyden Y (2004) LC-GC Europe, 17, 586
10. Massat DL, Vander Heyden Y (2004) LC-GC Europe, 18, 84
11. de Juan A, Casassas E, Tauler R (2000) Encyclopedia of analytical chemistry: intrumantation and applications. Wiley, New York
12. Findeisen W, Szymanowski J, Wierzbicki A (1980) Teoria i metody obliczeniowe optymalizacji. Państwowe Wydawnictwo Naukowe, Warszawa
13. Brandt S (1998) Analiza danych. Państwowe Wydawnictwo Naukowe, Warszawa
14. Acheson RM (1956) Acridines. Wiles, New York
15. Dzięgielewski J, Składanowski A, Konopa J (1996) Ann Oncol 7(Suppl., 82)
16. Berger B, Marquardt H, Westecdorf J (1996) Cancer Res 56:2094
17. Moriyama Y, Takeda K, Murakami K (2000) Langmuir 16:7629
18. Murakami K (1998) Bull Chem Soc Jpn 71:2293
19. Murakami K, Tsurufuji K (1999) Bull Chem Soc Jpn 72:653
20. Burger AM, Double JA, Konopa J, Bibby MC (1996) Br J Cancer 74:1369

Part II
New Techniques and Special Applications

Chapter 3
Surface-enhanced Raman Scattering (SERS) in Bioscience: A Review of Application

Jolanta Bukowska and Piotr Piotrowski

Abstract This article reviews some recent applications of surface-enhanced Raman scattering (SERS) spectroscopy in biology, biochemistry and biomedicine. We start with a short description of theoretical background of Raman scattering enhancement by the plasmonic nanoparticles. SERS is a phenomenon observed for molecules in a close proximity to the surface of metallic nanostructures. We present an overview of SERS substrates fabricated using various physical and chemical methods. SERS spectroscopy, which combines very high sensitivity with molecular specificity, is a powerful technique for studying biologically important systems, ranging from simple molecules like amino acids, to extremely complex samples such as living cells and tissues. We demonstrate great potential of SERS not only for detection and identification of (bio)molecules, but also in monitoring various biochemical processes. The strategies that are used for biosensing with the aid of SERS spectroscopy are briefly described.

Keywords Surface-enhanced Raman scattering (SERS) • SERS of biological molecules • Biosensors • Intracellular SERS • Nanoparticle probes

3.1 Introduction

Inelastic scattering of light by molecules, utilized by Raman spectroscopy, is nowadays commonly used to provide valuable information on molecular structure as well as chemical composition of the studied samples. Conventional (or normal) Raman spectroscopy suffers from very low scattering cross-sections compared with other spectroscopic methods. This significantly limits its applications to investigate molecules in solutions at low concentrations, which is a common requirement in analytical chemistry and biological systems. The simplest way to increase Raman intensity is exciting the spectrum with the laser beam of the energy corresponding to the energy difference between ground and excited electronic state. This effect, called resonance Raman (RR) scattering allows us to enhance the spectrum several orders of magnitude. However, in case of colorless samples, this method requires excitation in ultraviolet (UV), while the most popular light sources in Raman

J. Bukowska (✉) · P. Piotrowski
Department of Chemistry, University of Warsaw, Pasteur 1,
02-093, Warszawa, Poland
e-mail: jbukow@chem.uw.edu.pl

M. Baranska (ed.), *Optical Spectroscopy and Computational Methods in Biology and Medicine,* Challenges and Advances in Computational Chemistry and Physics 14,
DOI 10.1007/978-94-007-7832-0_3, © Springer Science+Business Media Dordrecht 2014

spectroscopy are visible or near infrared lasers. Another, more advanced possibility to increase the Raman scattering intensity is the use of non-linear effects such as stimulated Raman scattering. However, the most effective way of amplification of the Raman scattering is offered by surface enhanced Raman scattering (SERS) spectroscopy. In SERS, Raman signal may be enhanced by as much as 11 orders of magnitude. Typical SERS enhancement factors are of order of 10^4–10^6. Further enhancement of the Raman spectrum is obtained when the laser excitation energy corresponds to electronic transition of the molecule (resonance Raman). In this effect, called surface enhanced resonance Raman scattering (SERRS), enhancement factors can reach 10^{11}, allowing detection at a single molecule level. Combined enhancement processes provide a Raman signal of sensitivity at least equivalent or even higher than that of fluorescence. Very high sensitivity, selectivity and ability to identify molecular species, together with other advantages make SERS/SERRS spectroscopy an exceptionally attractive tool for variety of applications in bioscience. SERS may compete successfully with other spectroscopic methods in bioscience, such as fluorescence.

Giant amplification of the Raman signal in SERS is generated by metallic nanoparticles (typically silver, gold and copper), on which target molecules are adsorbed. The first surface-enhanced Raman spectrum of pyridine adsorbed on a silver electrode was observed by Fleischmann and coworkers in 1974 [1] at the University of Southampton, but in fact they did not recognize the importance of their results. The first explanation of this new phenomenon in terms of a giant surface enhancement of the Raman scattering, observed by Fleischmann et al., was reported by Jeanmaire and van Duyne in 1977 [2]. Since 1977, many papers on both theoretical and practical aspects of SERS spectroscopy have been published. In the last several years, popularity of surface-enhanced Raman spectroscopy has grown rapidly. Among a great variety of applications, these connected with biomolecule detection develop extremely quickly.

3.2 SERS Phenomenon—Background and Mechanism of Enhancement

Phenomena concerning interactions between matter and light might be roughly divided into three cases: absorption and emission of light, scattering—when, as opposed to emission, angle distribution of radiated photons is observed—and reflection, when the wave vector changes its direction at an interface between two media according to the laws of reflection. Raman effect falls into the second category and is defined more accurately as inelastic scattering of light, which means that the energy of the scattered photons differs from the incident ones. Thanks to this energy difference, Raman spectroscopy informs us about energies of rotational and vibrational modes of the studied molecule. Raman spectrum contains a set of bands, which is characteristic for a given molecule, thus identifies the sample, providing its so called chemical fingerprint.

Vibrational Raman effect is classically described by Placzek's theory, which is a simplified depiction of interaction between a molecule and electromagnetic wave that induces dipole moment in the molecule. As far as oscillation Raman effect is concerned, one needs to realize that molecular vibrations influence properties of the charge distribution in the molecule. Therefore, dipole oscillates with frequencies ν_0, $\nu_0 - \nu_{osc}$, and $\nu_0 + \nu_{osc}$, emitting light of frequency equal to its own (ν_0 being frequency of incident light beam and ν_{osc}: frequency of molecular vibrations).

As it follows from the classical description, two types of Raman scattering are distinguished: when the energy of the scattered photons is lower or higher than the energy of the incident ones. It is explained as losing or gaining energy to vibrational or rotational modes of the molecule.

Unfortunately, Raman effect concerns only a small number of photons; vast majority of incident photons is scattered elastically without the energy change. Raman scattering involves only 1 of about 10^7 photons. Such obstacle prevented this technique from effectively developing in application fields for years. It was only in 1970s when Fleischmann et al. [1] observed a sudden rise in the Raman scattering intensity of the molecules adsorbed on a roughened electrode surface. Even though the contribution of the metal surface to the enhancement was not initially recognized, this technique is fully referred to as surface-enhanced Raman scattering (SERS) nowadays. Intrinsically weak Raman scattering can be enhanced thanks to the interaction between all three components of the system: light, adsorbed molecule and the metal surface. Nature of the effect, which is a borderline case between chemistry and physics, makes it necessary to combine languages of both fields in order to fully describe its mechanism. Theories explaining SERS belong to one of the two categories: chemical and electromagnetic ones. It is now well known that the dominant contribution to the enhancement arises from the electromagnetic effect. However, some observations, such as stronger enhancement of the Raman scattering of chemically adsorbed molecules as compared to physically adsorbed ones cause that chemical mechanism should be also taken into account.

Chemical explanation of the surface-enhanced Raman scattering mechanism makes use of charge-transfer theory. It focuses on the influence of the metal surface on energy levels of adsorbed molecules [3]. According to the chemical theory, Raman intensity gain in SERS is obtained by matching the energy of the incident beam with the energy of electron transition between Fermi level of the metal and a molecular level of the adsorbate. Situation when intensity of Raman scattering rises due to laser-induced intermolecular transitions is analogous to resonance Raman mechanism. In that case, enhancement of the Raman scattering is obtained when a laser beam of energy equal to electronic transition of the molecule is applied.

Some features of SERS prove that chemical mechanism plays a role in the enhancement; potential-dependent experiments show that the phenomenon might be tuned to a certain beam energy through applying appropriate bias to the metal adsorbent. However, there are features that cannot be understood on the ground of charge-transfer theory. It turns out that every SERS-active metal has its own energy work range. It is widely known that silver nanostructures yield the best enhancement when exposed to blue/green laser beam; on the other hand, red light supports SERS

on gold structures. In addition, enhancement factor depends on electric permittivity of the medium. This inconsistency led to the second theory of SERS: electromagnetic one. It explores the influence of electromagnetic wave on a roughened surface. First attempt to describe SERS electromagnetically was dipole image theory. It presented a molecule as an oscillating electric dipole. Vibrations of the induced dipole are caused by oscillating electric compound of the electromagnetic wave. As the molecule is situated near the surface, it polarizes the metal. Thus, image dipole in the metal is created. It, in turn, causes extra polarization of the molecule which completes the feedback loop of interaction leading to the enhancement of the electric field near the metal surface. One of the main flaws of this theory is that it does not answer the question why metal nanostructures on the surface are essential to Raman enhancement. More refined theory includes the idea of surface plasmon resonance (SPR). Physicists define plasmons as collective oscillations of electron plasma (collection of valence electrons in the solid) [4]. Plasmons of different energies describe different excitation modes of such oscillations. Special properties of longitudinal charge density waves propagating at a metal/dielectric interface led to distinction of a subgroup of plasmons: surface plasmons [5]. These quasi-particles can be excited only by a transverse electromagnetic wave. What is more, optical properties of the system change drastically when energies of surface plasmons and incident light beam match. In that case, giant rise in electric field intensity near the interface is observed. This phenomenon is called surface plasmon resonance. For light from near UV/VIS/near IR range (usually used to excite Raman scattering), SPR is possible to obtain for plasmons confined in small metallic nanostructures; then we talk about localized surface plasmons (more accurately localized surface plasmon-polaritons, in order to refer to coupling between plasmons and light). When SPR occurs, molecules situated in the vicinity of such a metal nanostructure are exposed to electric field of intensity magnified several orders of magnitude (in comparison to the incident beam). As a result, electromagnetic enhancement factors reach up to 10^{11} (see critical paper on this topic by Ru et al. [6]). Exceptionally high enhancement factors are seen on the edges of the nanostructures. Such places might be called hot-spots (Fig. 3.1). Needless to say, only selected metals enhance Raman scattering effectively in the visible and near infrared range. Coinage metals, silver, gold and copper, face all the extra requirements needed to provide SERS signal. Extensive outlook on the problem of SERS substrates is presented in sect. 3.

One more thing needs to be highlighted. If one is already acquainted with Raman spectroscopy and knows Raman spectra of investigated molecules, it might be tempting to compare them to SERS spectra. However, they may—and usually they do—differ. Due to adsorption onto the metal surface, energies of molecular vibrations change which leads to modification of the Raman shift of certain bands. Furthermore, selection rules that apply to electromagnetic and charge transfer mechanisms of surface enhancement make some vibrational modes selectively enhanced, dependent on the relative contribution of each mechanism to total enhancement [8]. Thus, relative intensities of the bands in the SERS spectra are also altered in comparison to the normal Raman ones. As normal Raman selection rules do not apply in SERS, some bands ascribed to forbidden transitions may appear. Signifi-

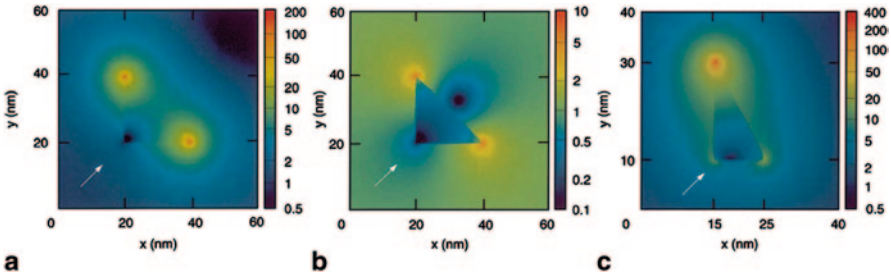


Fig. 3.1 Mathematic modelling of electric field amplitude distribution in the vicinity of different triangular silver nanoparticles as a result of surface plasmon resonance at different excitation wavelengths: (**a**) 412 nm, (**b**) 600 nm, (**c**) 458 nm. (Reproduced from Ref. [7] with kind permission of Elsevier BV)

cant rise in the background is usually observed in the SERS spectra. Among less apparent differences one may recognize: increase of the depolarization ratio of the bands (intensity ratio of the perpendicular to the parallel component of the Raman scattered light) and potential dependence of positions and intensity of the bands for molecules adsorbed on electrodes.

3.3 SERS Supports

For generating SERS spectrum, molecule must be in close contact with SERS-active support. Although many materials have been reported as providing enhancement of the Raman scattering with visible and near-infrared excitation, there are three metals that exhibit the highest enhancement—silver, gold, and copper. Because of high reactivity of Cu, silver and gold are most commonly used as the SERS supports. A key problem in a variety of applications of SERS is obtaining substrates that provide stable, robust, reproducible and effectively enhancing surface. Development of strongly enhancing and highly reproducible substrates is critical in the applications of SERS as an important and efficient spectroscopic technique for studying biomolecules. Some examples of SERS supports are shown in Fig. 3.2.

The main source of the surface enhancement is a very strong electromagnetic field generated by excitation of the localized surface plasmons in the metallic nanoparticles (NP) (see Sect. 2). It is well known that the highest enhancement of the Raman spectrum is obtained when the energy of the laser beam used to excite the spectrum is close to the energy of the surface plasmons. To fulfill the condition of the plasmon resonance for a given excitation laser line, nanoparticles should have proper size, shape and composition. Historically, first substrates used for SERS were electrochemically roughened Ag electrode surfaces [1, 2]. They are still used in some SERS experiments, especially when the applied potential has to be controlled. The advantage of the electrodes is an ease of electrochemical preparation of the rough surface, high enhancement factors and possibility to simply regenerate the surface by subsequent electrochemical dissolution—deposition of the electrode material.

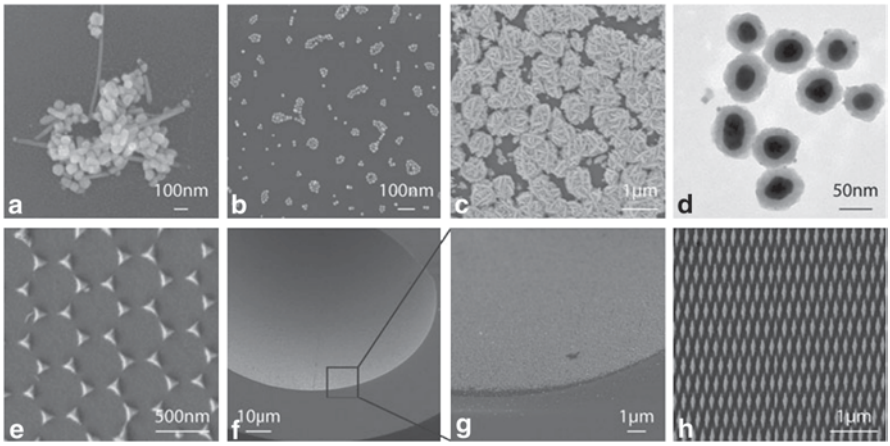


Fig. 3.2 Overview of various SERS substrates: scanning electron microscopy (SEM) (**a**–**c**, **e**–**h**) and transmission electron microscopy (TEM) (**d**) images of **a** aggregated silver nanoparticles produced by a modified Lee–Meisel protocol, **b** gold nanoparticles arranged in clusters, **c** flower-like silver composites prepared by an enzymatic silver deposition, **d** SERS tags for labeling in bioanalytics consisting of gold cores coated by a silica shell, **e** regular arranged metallic triangles by applying nanosphere lithography, **f** optical fibers as templates for SERS substrates, **g** deposited nanoparticles within the optical fiber (magnification of Fig. 3.2f), **h** electron beam lithographic top-down fabrication of periodically arranged gold nanorhombs. (Reproduced from Ref. [33] with kind permission of Springer-Verlag)

Soon after discovery of the SERS effect, metal (Ag and Au) colloids have been introduced as a very effective support (see Fig. 3.2) [9, 10]. Colloidal nanoparticles are produced mainly by chemical reduction of silver or gold cations in solutions, using reducing agents such as sodium citrate, sodium borohydride or hydroxylamine hydrochloride [11]. Metallic nanoparticles in colloidal suspensions are still the most widely used SERS active substrates, being ideal and simply produced substrates for performing measurements in solution. Main drawback of the chemically produced colloidal suspensions is some irreproducibility of the enhancement factors, caused by aggregation processes [12]. To overcome this problem, various methods of fabrication of metal nanoparticles have been developed. For some applications, including those in biology and biomedicine, metal nanoparticles obtained in colloidal suspension are deposited on the solid substrates. However, when colloidal particles isolated on planar substrates were used as the enhancing medium for ultra-sensitive detection of adenine, additional anomalous bands in the background were observed. Significant reduction of the spurious signals was obtained for hydroxylamine-reduced colloid, as compared to the more common citrate- and borohydride-reduced silver colloids [13].

There are many strategies of NP immobilization on solid supports, among which chemical or electrostatic self-assembling are the most popular. Chemical methods consist of functionalization of the solid substrate (usually glass or quartz) with thiol (SH) or amine (NH_2) groups using bifunctional molecules which are able to attach

Ag or Au nanoparticles from the colloid suspension. This strategy was first demonstrated by Chumanov et al. in 1995 [14] and developed by group of Natan et al. [15]. The idea of depositing metal nanoparticles from colloidal suspensions was also successfully applied to the fabrication of three-dimensional (3-D) structures of NP at the solid substrates by repeating chemical binding of the subsequent layers of NP deposit [16–18]. Another way is immobilization of NP by electrostatic attraction of the charged metal particles by oppositely charged solid support. In this strategy the most popular method is covering the solid substrate with positively charged derivatized silanes (e.g. aminopropyltrimethoxysilane) [14] or polymer layer such as poly(diallydimethylammonium chloride), which, as a positively charged polymer, is able to attract the negatively charged colloids [19] or polylysine [20]. Another, less common substrates include surfaces roughened by chemical treatment of solid substrate [21], "island films" deposited on glass and films deposited by evaporation or sputtering substrates [22, 23].

Metallic nanoparticles or other nanostructures can be also fabricated directly on the surface of the solid support using advanced physical methods such as nanolithography [24–28] or micro-contact printing [29]. Thin films of silver or gold can be also vapor-deposited over polystyrene or silica nanospheres of controlled diameter (FON-film over nanospheres), deposited on solid substrate. This method results in highly reproducible and effective supports [30, 31]. Using this technique, van Duyne group developed interesting SERS support by removing the underlying nanosphere mask, thus leaving periodic structure of metal nanoparticles (Fig. 3.2e), deposited on the free space between the nanospheres. As reported [32], supports prepared in this way exhibit very high enhancement factors (of order 10^8).

Plasmonic properties of metal nanoparticles can be significantly changed by tailoring their shape or size. Numerous methods of production of shaped nanoparticles such as nanorods [34, 35] triangles [36], nanocubes [37, 38], polyhedrons [39], stars [40–42] and nanoflowers [43] have been developed. Some of these nanostructures exhibit higher enhancement factors as compared to nanospheres, because of the presence of sharp edges and corners that are able to concentrate extremely strong electromagnetic fields (see Fig. 3.1). Among a great variety of nanoparticles, hollow spheres with differing cavity diameter—which allows tuning the plasmon properties—are very effective enhancers [44].

The possibility of shifting plasmon resonance to the near infrared is sometimes important in biological applications, because of problems with surface photochemistry when using visible radiation (especially on silver) and/or problems with fluorescence background. This red-shift may be also achieved in core-shell nanoparticles built of a spherical silica core of varying dimensions, covered with a thin metal shell [45, 46]. It was confirmed that spherical Au nanoshells, in which plasmon resonance frequencies are controlled by the relative inner and outer radius of the metallic shell layer (see Fig. 3.3), can be used as an effective and reproducible SERS substrate, also for large biomolecules and more complex systems such as DNA and living cells [47–50].

Interesting SERS substrate has been developed by Tian group [52, 53]. They coated Au nanoparticles with a very thin (2–3 nm) silica film that isolates the metal

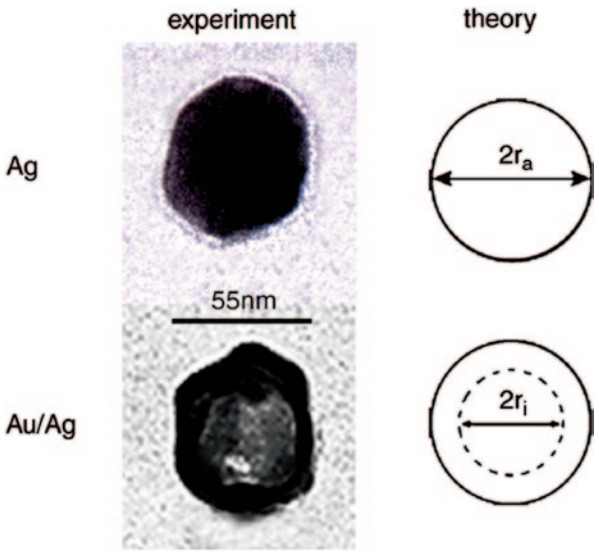


Fig. 3.3 (*Left*) TEM images of a Ag nanoparticle (*top*) and a hollow Au/Ag nanoshell (*bottom*). (*Right*) Schematic diagram of the particle geometry. (Reproduced from Ref. [51] with kind permission of Elsevier Science)

from the environment, preserving high enhancement factor of the shell-isolated nanoparticles (SHINERS—shell-isolated nanoparticle enhanced Raman spectroscopy). Protective layers of silica [54], glass [55] or polyethylene glycol [56] are also applied for metallic nanoparticles labeled with various compounds called Raman reporters. These encapsulated nanoprobes find many applications as SERS probes incorporated in a variety of biological materials (see Sect. 4.2). Larger enhancement of the Raman scattering than that produced by a single nanoparticle (NP) is created at the junction between two nanoparticles in a dimer, in which nanoparticles are separated by nanometer-scale gaps [57–60]. Such NP dimers are sometimes called hot spots. Reproducible fabrication of these nanostructures is however challenging, because the SERS enhancement factor critically depends on interparticle spacing [61].

This short review does not cover all developed SERS supports. Usually, SERS substrates are prepared in a given laboratory, but effectively enhancing gold solid support known as Klarite, or colloid suspensions are now also commercially available.

Another approach to enhance Raman scattering is offered by tip-enhanced Raman spectroscopy (TERS), reported for the first time in 2000 [62] and significantly developed in the last ten years [63–67]. In TERS, very strong electromagnetic field is located at the nanometer-scale metallic tip (Ag or Au). In an ideal case, it ends with a single metal nanoparticle. Its radius is much smaller than the diffraction limit so it allows measuring the spectrum from a much more confined spot. The Raman signal is strongly enhanced when tip is brought close to the sample irradiated by the laser beam (the movement is controlled by atomic force or scanning tunneling mi-

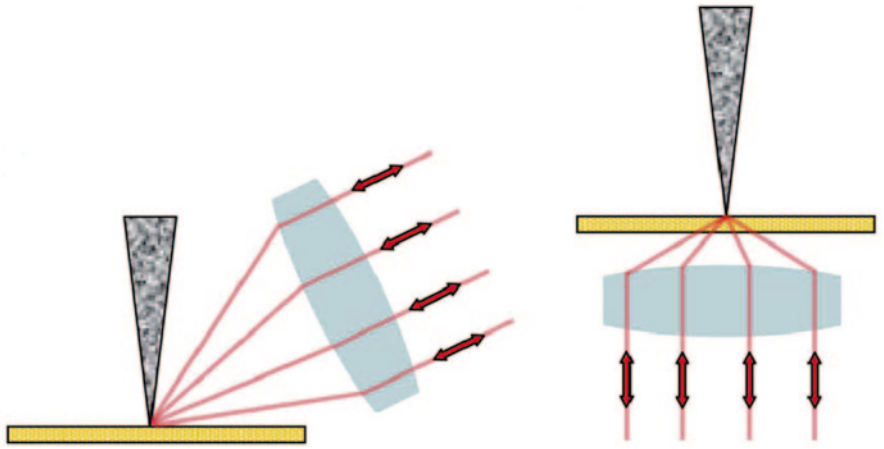


Fig. 3.4 Different configurations used in TERS experiments; *left*—side illumination, *right—bottom* illumination, using an inverted microscope. (Reproduced from Ref. [67] with kind permission of Elsevier BV)

croscopy head) (Fig. 3.4). Therefore, TERS provides spectroscopic characterization of the sample with a nanometric spatial resolution.

To date, many TERS experiments have been reported for a variety of chemical and biochemical systems. They have been thoroughly reviewed in 2008 by Bailo and Deckert [68]. In the last ten years, TERS spectroscopy has been significantly developed and opened new possibilities of applications in a variety of fields in biology and medicine. In particular, experiments with living cells seem to be very promising. Interesting review that shows the potential and challenges of TERS applications in bioscience was published in 2010 by Deckert-Gaudig and Deckert [69].

3.4 SERS in Biology, Biochemistry and Biomedicine

3.4.1 SERS of Aminoacids, Peptides, Proteins, and Enzymes

Applications of SERS in bioscience cover very broad area, from relatively simple systems such as aminoacids and peptides through proteins, nucleic acids and enzymes, up to individual cells, living tissues, bacteria and viruses. Advantages of SERS spectroscopy in structural studies of small, biologically important molecules such as amino acids, purine and pyrimidine bases, porphyrins, flavines etc. have been shown just after discovery of the SERS phenomenon. These early studies have been reviewed by Cotton in 1988 [70]. SERS of amino acids and proteins has been observed on various SERS active substrates, among which silver and gold colloids are the most important and the most popular. However, it is well known that

SERS spectra on metal colloids suffer from some irreproducibility [71–73] caused by aggregation of metallic nanoparticles. On the other hand, aggregating agents are indispensable to achieve better surface enhancement, because huge electromagnetic fields responsible for surface enhancement are created at particle junctions (nanogaps). A common problem in these experiments is the inhomogeneity of the colloids, which in turn provides local differences in the SERS enhancement factors. Most SERS studies of amino acids aimed at establishing the mode of interaction with the metal and the geometry of the molecule at the surface. SERS spectra allow determining the molecular form of amino acid present at the metal support (cationic, zwitterionic, or anionic), owing to high sensitivity of the spectrum to the molecular structure of the adsorbed species [74–79]. Some experimental SERS results are supported by theoretical calculations using DFT approach, making interpretation of the spectra more credible [80].

The most stable and reproducible SERS spectra of aminoacids have been recorded for sulphur containing molecules such as cysteine or methionine [77, 81–83]. SERS spectra of cysteine were recorded on various substrates such as colloidal nanoparticles [84] and rough silver electrodes, on which cysteine spontaneously self-assembles [81, 83]. Cysteine molecules are chiral—they can exist in one of two enantiomeric forms (L- or D-), that only differ in that their structures are mirror images of each other. Because only one enantiomer tends to be physiologically active while the other one is inactive or even toxic, drug compounds are produced in an enantiomerically pure form. Chiral surfaces may be applied as sensors for chiral recognition. In principle, both: the normal Raman and the SERS spectra are not sensitive to molecule chirality. However, in carefully designed experiments SERS spectra of cysteine monolayers on the Ag electrode surfaces have been shown to be sensitive to chirality of adjacent cysteine molecules at the metal surface [83]. It was demonstrated that there is a range of electrode potentials in which cysteine molecules are adsorbed mainly in zwitterionic form with the COO– groups close to the surface, while at more negative potentials NH_3^+ groups deprotonate at the surface with simultaneous weakening of the interaction of the carboxylic groups with the surface. As expected, the potential-induced effects for D-cysteine were similar to these for L-cysteine. However, for racemic mixture at acidic pH, the changes in the spectral pattern corresponding to potential-induced transition from adsorbed zwitterions to neutral molecule were considerably smaller. The origin of the observed effects were explained in terms of the most stable adsorption configuration for cysteine that involves metal-sulphur, metal-nitrogen and two hydrogen bonds between carboxylic groups (cyclic dimer) as indicated by DFT calculations [85]. In the case of LD cysteine dimers, there is a mismatch in the carboxylic bonds of the neighbouring molecules, which results in considerable weakening of the hydrogen bonds between them. The L- and D- cysteine enantiomers adsorb at the metal surface with equal probability, when adlayer is self-assembled from the solution. Thus, in spite of higher stability of the homochiral dimers, molecules of opposite chirality may exist at the surface in the immediate vicinity. In such a case intermolecular interactions between adjacent molecules are considerably weaker than for monolayer formed by the molecules of the same chirality, what is reflected in the band positions in the respective

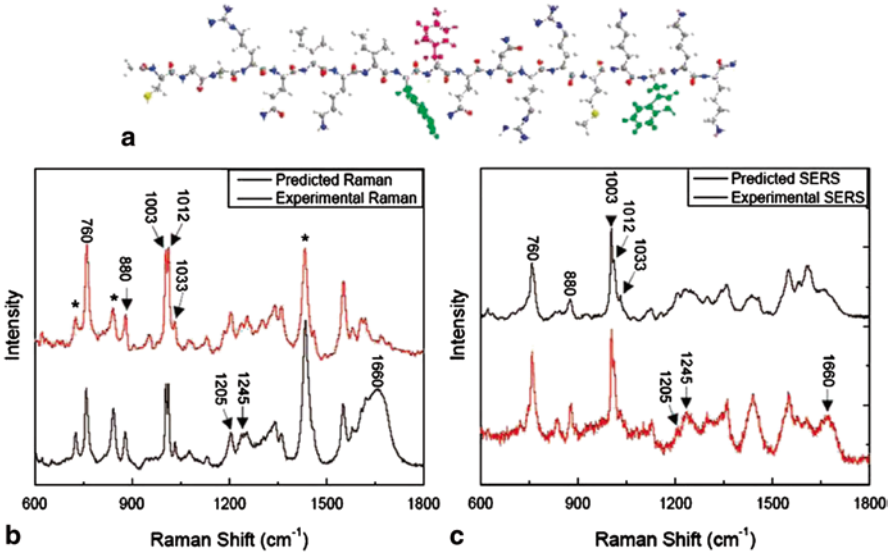


Fig. 3.5 Comparison of empirically predicted (*red*) and measured (*black*) spectra of penetratin. (**a**) Molecular model of penetratin, including one phenylalanine (*purple*) and two tryptophan (*green*). (**b**) Raman spectra (peaks assigned to tryptophan are denoted with an asterisk). (**c**) SERS spectra. (Reproduced from Ref. [91] with kind permission of the American Chemical Society. Copyright 2008, American Chemical Society)

SERS spectra. Another example that confirms sensitivity of the SERS spectra to chirality is recognition of phenylalanine enantiomers by the enantiomeric cysteine monolayer deposited on a silver support [86] and SERS evidence for chiral discrimination in intermolecular interactions between adsorbed methionine molecules [87]. Further exploration of the stereospecificity of SERS may open new possibilities of applications of SERS spectroscopy in optical sensors of chiral biomolecules.

TERS is a very helpful technique in investigating molecules of biological importance. In 2009, first TERS spectra of cystine adsorbed on ultrasmooth gold substrate were reported. They indicated local variations of cystine adsorption [88]. It was also possible to observe two differently protonated histidine forms and determine two different geometries of adsorption on an atomically flat silver substrate [89]. SERS data for simple amino acids made it possible to interpret the SERS spectra of more complex systems such as peptides [90–95], dipeptides [96, 97], and proteins [98–100]. For example, it was possible to predict major features of the SERS spectra of penetratin, a 19 amino acid cell penetrating peptide, based on the SERS spectra of aromatic aminoacids such as tryptophan and phenylalanine and cysteine-containing aromatic peptides bound to Au nanoshells [91]. As may be seen in Fig. 3.5, the SERS spectrum of penetratin is dominated by features characteristic of the aromatic amino acids, greatly simplifying its interpretation.

While interpreting complex SERS spectra of peptides and proteins, it must be kept in mind that the most strongly enhanced bands in the spectrum correspond to the vibrational modes involving part of the molecule in the vicinity of the metal

support. It has been proved that electromagnetic enhancement decays with distance from the surface, but is still measurable at about 10 nm [101]. Therefore, in case of very large biomolecules, spectral information may be confined to some fragments of the molecule directly interacting with the metal nanoparticle or in a short distance from the metal surface. This specificity and selectivity of the SERS spectra cause that in the case of large bio-molecules some fragments of the molecules are not "seen" in the spectra. On the other hand, this would enable characterization of the interaction between metal surfaces and biological molecule and determination of its orientation at the surface.

Also TERS spectra of peptides have been successfully recorded. TERS studies of peptide adsorption on metal surface (gold nanoplate) were carried out on a model molecule of oxidized glutathione and the structure of the peptide on the Au surface was determined [102]. The uniformity of the spectra measured at several points along a line across the gold substrate indicated the consistent orientation of the peptide. This result was regarded as crucial for possibility of characterizing and sequencing peptides with the help of this method.

Another problem related to SERS biosensing of proteins is their denaturation upon contact with metal nanoparticles. Thus the biocompatibility of the metal surface must be carefully controlled. The most simple way is to cover the nanoparticle surface with a self-assembled monolayer (SAM) [103] that prevents direct contact between the protein and the metal surface. There is a great variety of SAMs that could be used to prevent denaturation of the protein at the metal surface. However, the most optimal ones have to effectively bind the protein e.g. through electrostatic or covalent interaction, with simultaneous preservation of the protein properties and biological activity. The most popular SAMs are built of ω-substituted alkanethiols $HS(CH_2)_nX$, which form very stable metal-sulfur chemical bond. They create an interface with a well-defined composition and structure and with easily controlled properties achieved through the selective use of X groups. Among alkanethiols with negatively charged terminal groups, carboxylate-functionalized ones are the most frequently employed [104–109]. Sulfonate group is an example of an anionic group that can be used alternatively [110]. A great deal of work has been devoted to design biocompatible surfaces that ensure preserving the native structure of immobilized proteins. In the case of redox protein, communication between protein and electrode, which enables effective electron transfer, is also a major challenge. The case of cytochrome c (Cc) is a good illustration of the advantages for using SERS in studies of redox proteins. Numerous reports have been published on SERS of Cc immobilized on a silver or gold support modified with various monolayers [105, 108–112]. To improve intensity of the spectra, they are recorded using excitation laser line in resonance with molecular transition in Cc chromophore (SERRS spectra). The SERRS spectrum of Cc not only allows identifying the protein immobilized on the surface, but also makes it easy to determine the oxidation state and spin configuration of iron in heme groups, because Raman bands are well known markers of these properties [113, 114]. Moreover, sensitivity of the SERS spectrum to the orientation of the adsorbed molecules with respect to the surface causes that relative orientation of the heme plane may be determined by using so called surface

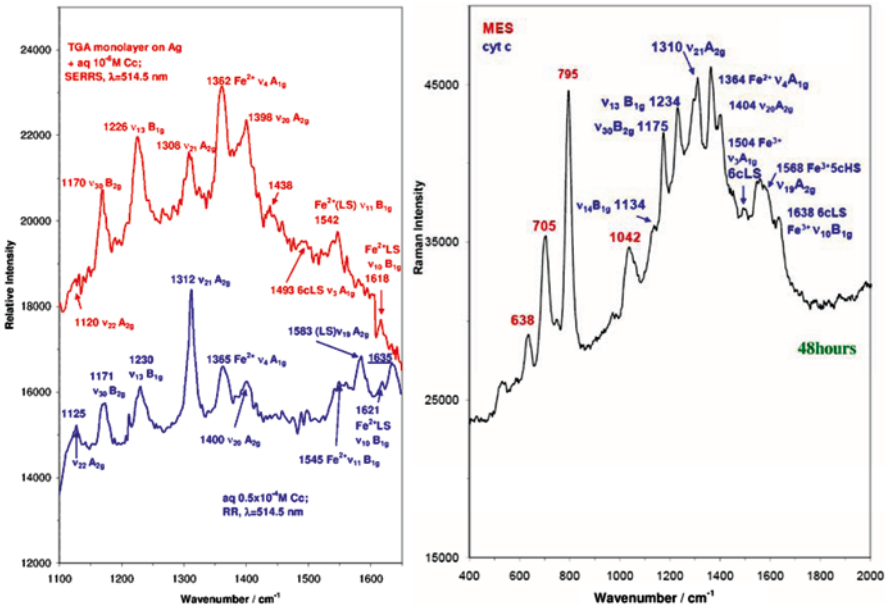


Fig. 3.6 SERRS spectra of cytochrome c adsorbed from 10^{-4} M aqueous solution of Cc on Ag electrode, coated with monolayer of thioglicolic acid (TGA) (*upper* spectrum—*left* panel) (A. Królikowska, unpublished results) and mercaptoethanesulphonate (MES) monolayer (*right* panel), as compared with the resonance Raman spectrum of 0.5×10^{-4} M Cc solution. (Reproduced from Ref. [110] with kind permission of John Wiley and Sons)

section rules [8, 115]. Figure 3.6 shows SERRS spectrum of Cc electrostatically immobilized on silver electrode, which has been covered with a monolayer of thioglicolic acid (TGA) and resonance Raman spectrum of Cc solution. As may be seen, both spectra exhibit different frequencies and intensity pattern of the bands, which excludes risk that the bands detected in the SERRS spectrum originate from the protein in the solution. The band positions in the SERRS spectra are characteristic of the native 6cLS configuration of heme iron. Namely, the band at 1362 cm^{-1} indicates Fe^{2+} oxidation state, together with the 1493, 1542 and 1618 cm^{-1} bands which are markers of the native low spin (LS) configuration. On the contrary, in some cases bands assigned to oxidized form of Cc (at 1568 and 1638 cm^{-1}) are clearly visible as for Ag support modified with mercaptoethanesulphonate (MES) (see spectrum in Fig. 3.6). The band at 1568 cm^{-1} indicates lots of Met-80 axial ligand leading to a non-native B2 5cHS state of heme and 1638 cm^{-1} feature is ascribed to the native B1 6cLS state [110].

Potential-dependent SERRS studies for Cc immobilized on functionalized SAMs on Ag allow monitoring conformational changes in interfacial redox processes [111]. SERRS experiments may be performed in the stationary conditions with continuously changing electrode potential [111, 116], as well as in time-resolved mode by applying a potential jump [117]. These spectroelectrochemical ex-

periments indicated that for the short thiol linkers the electron-transfer kinetics is determined by protein dynamics rather than by electron tunneling probabilities and that the reorientation of protein is controlled by the interfacial electric field. Thus, they show that SERRS can provide a very deep insight into dynamics of interfacial processes of proteins on the molecular level. SERRS experiments deliver much information about the heme part of the protein. A unique possibility of observing vibrational spectrum of both the heme and amino acid residues is however offered by TERS spectroscopy as demonstrated in 2008 [118].

Surface enhanced Raman scattering spectroscopy is now widely reported for enzymes. In this case, combination of SERS and resonance Raman spectroscopy is also usually applied, because it offers considerably higher sensitivity as compared to SERS. Excellent review on the development of SERS for the detection of enzyme transformations was published by Larmour et al. in 2010 [119]. Initially, SERRS was used for direct detection of enzymes. First results on SERRS of glucose oxidase has been reported at the very beginning of the SERS studies [120]. The first studies, in which biological activity of enzyme immobilized on Ag nanoparticles (Ag colloid) was monitored by SERRS spectroscopy, were published in 1993 [121]. In this report, 60–85 % retention of enzymatic activity of chlorocatechol dioxygenase in the reaction of catechol substrate with oxygen was demonstrated. The subsequent experiments for enzyme-substrate, enzyme-inhibitor, and enzyme-product complexes indicated that in general, SERRS allows probing the enzymatic processes *in situ*, during catalytic turnover. However, it has to be stressed that in all SERRS studies reported, some loss of the enzyme activity was observed upon contact with the metal support responsible for signal enhancement. Therefore, preparation of biocompatible substrates seems to be a key problem and important challenge in these experiments.

An alternative approach to study enzyme bioactivity with SERRS is recording the spectrum of enzymatically produced dyes, which are easily detected with SERRS. For example, azo dyes may be used in such experiments because they exhibit strong and very characteristic SERRS spectrum. SERRS signal of azoaniline adsorbed on Au nanoparticles has been applied for determination of glucose concentration [122], or adsorbed on Ag colloids for detection of antigen (mouse IgG) [123]. In this case azo dye was produced in peroxidase catalyzed oxidation reaction of o-phenylenediamine by hydrogen peroxide (Fig. 3.7).

The idea of utilizing SERRS spectrum of enzymatically produced dyes to detect enzymes was further developed by Stevenson et al. [124]. In a very interesting experiment, the SERRS spectrum of oxidized form of 2,2′-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS), used as a substrate in an enzyme-linked immunosorbent assay (ELISA), was applied for the indirect detection of human prostate specific antigen (PSA) [124]. ABTS yields a green product upon enzymatic reaction with horseradish peroxidase and is commonly used as a colorimetric indicator of peroxidase activity. Experiments by Stevenson et al. [124] proved that SERRS provides a lower limit of detection and a wider range of linearity of the Raman signal vs. concentration as compared to the standard colorimetric ELISA approach.

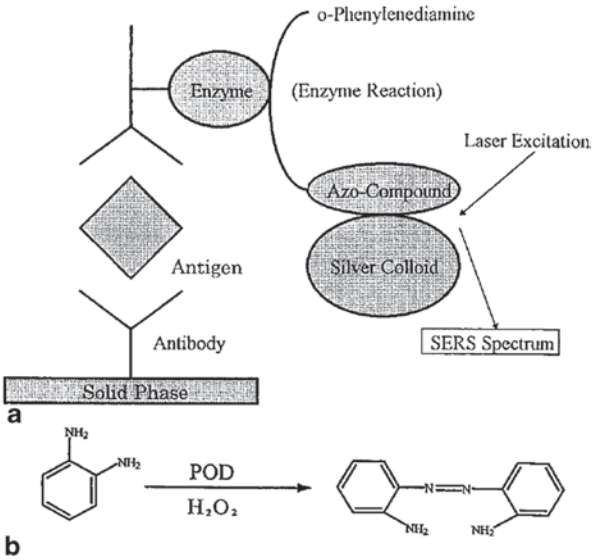


Fig. 3.7 SERS enzyme immunoassay system (**a**) and enzyme catalyzed reaction (**b**). (Reproduced from Ref. [123] with kind permission of the American Chemical Society)

Another example of successful application of SERRS for detecting the presence and measuring activity of the enzyme is the work of Ruan et al. [125], in which SERRS spectrum of a blue indigo dye produced by oxidation and dimerization of 5-bromo-4-chloro-3-indolyl phosphate was utilized for monitoring activity of alkaline phosphatase.

Detection of DNA is one of the most important tasks in bioanalysis. To date, numerous SERS measurements of DNA utilizing different SERS supports have been reported. Experiments can be roughly divided into two groups. In the first group, called label-free method, DNA is directly detected with the SERS nanoprobes. For example, Halas group reported SERS spectra of DNA oligonucleotides and found that they are selectively dominated by the vibrational modes of adenine [126]. The authors also demonstrated the capability of SERS in investigating the interaction of DNA with various molecules (in this case—cisplatin). While direct SERS measurements for DNA are rather scarce, there are many reports on the indirect SERS experiments. This approach to DNA detection consists of monitoring the SERS spectrum of reporter molecules (usually dyes) that are attached to the metallic nanoparticles. This kind of SERS experiments for DNA is more extensively described in sect. 4.2.1.

3.4.2 SERS Biosensors

Term "biosensor" has no clear definition. Therefore, it might be applied to many different systems. Generally, biosensors provide data about the biological analyte they are exposed to, monitoring a change in their physicochemical response. If we monitor a SERS spectrum coming from the sensor in varying conditions, we deal with a SERS sensor. There is a whole range of tasks which biosensors may be employed to: monitoring a chosen kind of molecules, identification of the analyte composition, or determination of the environment parameters. Previous section of this review describes SERS experiments used for detection of certain biomolecules such as proteins or enzymes. Those examples perfectly fit in what we call a biosensor—they give information about the presence of a chosen biochemical compound in the sample. Those SERS sensors make use of various types of metallic substrates to enhance Raman scattering. However, lots of opportunities in the cell analysis open if metallic nanoparticles are engaged: thanks to their size, they can be successfully introduced into the cell in order to report the intracellular conditions from a precisely defined place.

3.4.2.1 Specific Sensing of Biomolecules

Previously described experiments intended to detect biomolecules were carried out in the following manner: object of the study needs to be found in the close proximity of the Raman enhancing substrate and then its characteristic SERS spectrum is collected. There are cases, however, when such a direct approach cannot be applied. Problem of glucose sensing illustrates such situation perfectly. Glucose is a molecule of great importance in human body, especially due to its relation to diabetes which has been considered a disease of civilization in the Western World for years while its incidence in developing countries has been increasing rapidly. Monitoring a concentration of glucose with SERS is hindered by its small Raman cross section and minimal adsorption on SERS-active metals. Van Duyne et al. performed a series of experiments to improve and optimize sensing parameters of glucose [127–130]. They increased its affinity to the SERS substrate by adsorbing a SAM of decanethiol on the AgFON surface [127]. It let them detect glucose over a clinically interesting concentration range. Further investigations focused both on the composition of the partition layer and the substrate. Researches showed that the sensor with (1-mercaptoundeca-11-yl)tri(ethylene glycol) as a partition layer works also in presence of interfering specimens, such as bovine serum albumin [128]. Real-time sensing was possible with a mixed SAM consisting of decanethiol and mercaptohexanol [129]. Such a device exhibited stability for 10 days and was used to determine concentration of glucose in the environment of bovine plasma. On the other hand, it was proved that replacing AgFON with AuFON and 1-mercaptoundeca-11-yl)tri(ethylene glycol) with 1-mercaptoocta-8-yltri(ethylene glycol) resulted in red-shift of plasmonic resonance which could be a step forward towards apply-

ing NIR lasers in this area. It could reduce fluorescence and background spectra of biological matrix [130].

Another example of sensing biomolecules are SERS experiments with bacterial spores. Indirect SERS detection of some bacterial spores is based on tracking down its biomarkers—molecules which are released from the spores while absent from vegetative cells. In case of bacterial genus of *Bacillus*, dipicolinic acid (DPA) serves as a biomarker for its spores. Sensitive DPA sensor is interesting from medical point of view, as one member of the genus, *Bacillus anthracis*, causes an acute disease: anthrax. AgFON substrate turned out to be applicable in this problem as well [131]. Obtained limit of detection for *Bacillus subtilis* spores (~2600 spores per 0,2 µl) was well below anthrax infectious dose of ~10000 spores per 0,2 µl. Similar experiment was carried out using Au-nanoparticles-based substrate [132].

Various fields of biosensing make use of so called nanotags: nanoparticles providing Raman enhancement for molecules adsorbed on them. Having labelled a biomolecule of interest with such a nanotag, one may be able to detect it indirectly: SERS spectrum of a molecular dye adsorbed on a tag indicates presence of the studied analyte.

Adenosine triphosphate (ATP) biosensor based on such nanolabels was demonstrated by Li et al. [133]. Nanotag consisted of malachite green isothiocyanate (MGITC) molecules attached to gold nanostars which were additionally covered with a layer of silica to prevent the reporter from detachment. It was functionalized with a single-stranded DNA creating the signalling probe, while the capture probe: complementary DNA strand was immobilized on a gold substrate. Under normal conditions, aptamer with the nanotag attached and the complementary strand create a duplex DNA, thus immobilizing the tag on the gold surface. However, DNA dissociates in the presence of ATP which leads to decrease in the SERS signal from the nanotag. Such a device determines not only absence or presence of ATP but also its concentration.

DNA hybridization is also exploited in detecting DNA itself (Fig. 3.8). Dye molecule might be adsorbed on a SERS inactive substrate; collecting its SERS spectrum becomes possible in the presence of a target strand after its hybridization with a capture strand on the surface and a DNA-probe functionalized with a silver nanoparticle [134].

On the other hand, reporter molecule can be adsorbed on a metal nanoparticle bound to the probe DNA strand; it is only after hybridization of the probe strand with an appropriate target strand on the sample surface when collecting the SERS spectrum of the dye is feasible [135]. Detected DNA strand might be also labeled directly with a reporter molecule whose SERRS is collected, thanks to creating a duplex DNA with a complementary strand on a SERS-active surface [136]. By modifying different strands with different dyes, multiplexed DNA detection was also obtained [137].

SERS nanotags are widely used in immunoassays [138–140]. In this case, nanoparticles are functionalized not only with a reporter molecule but also with an antibody which introduces selectivity for a certain antigen. Antigen binds to the capture antibody immobilized on the surface and then to the antibody on the nano-

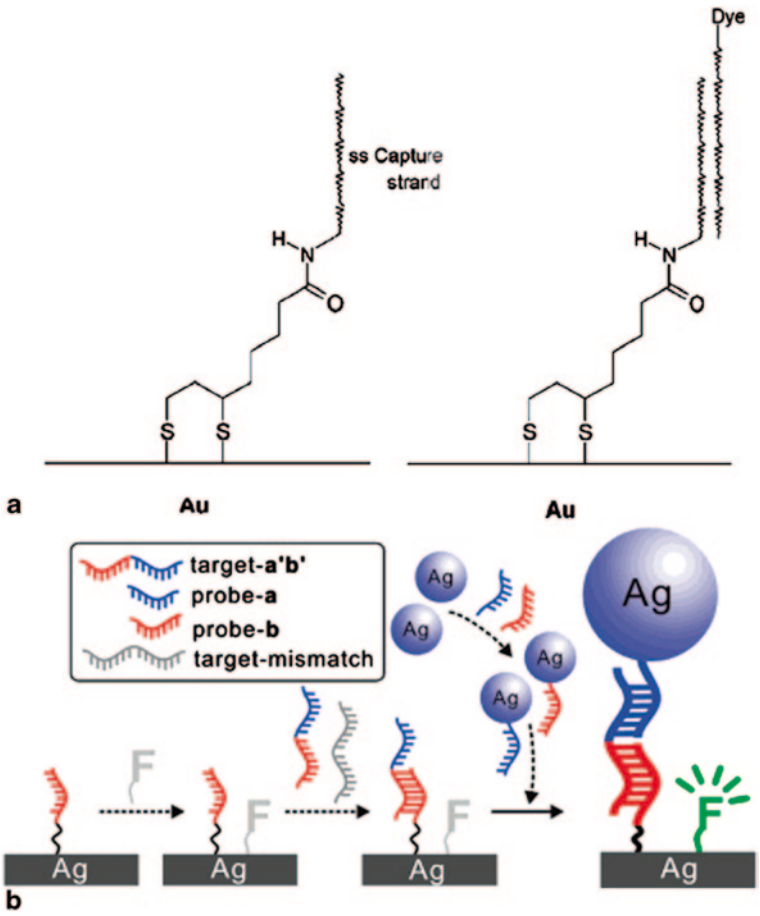


Fig. 3.8 Examples of different approaches in DNA SERS sensing: **a** SERS label bound directly to the DNA strand [136], **b** metal nanoparticle bound to DNA single strand while the label is adsorbed on the surface [134]. (Adapted from Refs. [134, 136] with kind permission of respectively The American Chemical Society, Copyright 2007 and Royal Society of Chemistry)

probe modified with the SERS reporter. Low limits of detection obtained in this approach make it competitive with other viral assays. It is possible to create a multi-sensitive assay by modifying nanotags for different antigens with different reporter molecules [141].

3.4.2.2 SERS Probing of Cells and Microorganisms

Very high sensitivity of SERS makes it an attractive method to study extremely complex systems such as single living cells. Compared to normal Raman spectra

of the cells, SERS offers considerably shorter collection times, which is especially important in Raman mapping. Application of SERS for probing living cells was reviewed in several papers [142–145]. First SERS report, which demonstrated the capability of SERS for cell studies was published in 1991 [146], but broader interest of scientific society to this subject actually started in 2005, as shown in an interesting review by Vitol et al. [145]. In SERS experiments with cells, nanoprobes that are responsible for enhancement of the Raman signal are inserted into the cells. Therefore, the key problem in intracellular SERS is development of efficient and biocompatible metal nanoprobes that do not affect cell integrity and viability. Nanostructures used in these experiments should be small enough to be transferred through the cell membranes without damaging them. Various SERS probes have been used for intracellular experiments, starting from the most popular and easily fabricated colloidal silver and gold nanoparticles [147, 148], through nanoshells [49], nanorods [149], nanostars [150], nanoflowers [151], to more sophisticated probes such as nanopipettes [152].

Generally, there are two ways of performing intracellular SERS experiments. In the majority of experiments, metallic nanoprobes are labeled with molecules that exhibit strong SERS (or SERRS) spectrum. They are called reporter molecules. Then, the SERS spectrum of the reporter molecule (e.g. a dye), linked to the nanoprobe is recorded. In the second, label-free approach, metallic nanoprobes enhance Raman spectrum of native constituents of the cell such as DNA, RNA, lipids or aminoacids, delivering chemical structural information at the molecular level. Coupling of the modern Raman instruments with confocal microscope opened a new possibility of obtaining Raman maps that show spatial distribution of investigated species with high lateral resolution.

In case of using label-free SERS nanoprobes the signatures of biological molecules and structures constituting the cell in the immediate nanoparticle vicinity can be delivered. In SERS experiments, nanoparticles are either introduced inside the cell or located on the cell membrane surface. First important report demonstrating power of SERS in such experiments was published in 2002 by Kneipp et al. [153]. It was possible to observe vibrational bands that were ascribed to amino acids, DNA and RNA and to find their distribution inside the intestinal epithelial cells. In 2006, Kneipp et al. reported SERS spectra from endosomes in living individual epithelial and macrophage cells at different times after the uptake of gold nanoparticles. [147]. Metal particles were found to be collected in endosomes inside the cells. Spectra indicated that the molecular composition in the vicinity of the nanoprobes changes over time. Spectra also varied between different cell lines. Moreover, total SERS intensity increased due to formation of nanoaggregates of which dimers and trimers were found to exhibit the highest enhancement factors. On the other hand, it is known that the increasing distance between aggregated nanoparticles results in lowering of the electromagnetic enhancement factor, that in turn result in decreasing intensity of the spectra. Uncontrollable aggregation of nanoparticles with time is a major obstacle in using metal colloids for SERS inside the cells, because of undesirable processes that result in changes of both the spectral pattern and SERS intensity [147].

SERS is also applied to microorganisms such as bacteria. It is difficult to introduce metal nanoparticles into bacterial cells because of their size, which is much smaller in comparison to mammalian cells. Therefore, the studies of bacteria with SERS are mainly focused on the extracellular experiments. However, Jarvis et al. reported intracellular SERS of *Geobacter sulfurreducens* using colloidal metal nanoparticles prepared from the enzymatic reduction of silver and gold salts by these bacteria [154]. Very interesting and critical review of the problems that are connected with the interpretation of the SERS spectra of such a complex molecular systems was published by Efrima and Zeiri [155]. They paid attention to considerable differences between the SERS spectra reported for the same bacterium by various authors. Efrima and Zeiri suggested that the differences observed are due to differences in preparation procedures that may result in interactions of nanoparticles with various chemical constituents of the cells. They demonstrated clear dependence of the spectra on localization of the nanoparticles. Unusual selectivity and sensitivity of SERS spectrum toward flavins, an essential coenzyme in cells, was observed for external (on the cell walls) coating of the bacteria with the metal particles. On the contrary, when the metal colloids were produced inside the cell, flavin signatures disappeared and the bands corresponding to internal chemical constituents appeared instead. To overcome the limitations of colloidal nanoparticles for SERS of the cells and microorganisms and to gain more control upon the position of the sensor, new approaches to intracellular sensing have been proposed. The most important are: using metal-coated fiber-optic tips [156, 157], SERS-active glass nanopipettes [145, 152] and tip-enhanced Raman scattering (TERS) [158, 159].

TERS technique, briefly described in Sect. 3, is a perfect tool in biosensing. Cell walls were successfully investigated with TERS. Thanks to sub-optical spatial resolution of this technique, it is possible to distinguish between protein and lipid components of the *Staphylococcus epidermidis* cell wall [158, 160]. Fluctuations in the spectra collected from one spot of the outer layer of the cell may be ascribed to the bacterial surface dynamics [161]. Surface of a single tobacco mosaic virus was also characterized with TERS. Not only protein components from the capsid, but also some RNA features were seen in the spectrum (Fig. 3.9) [159]. These very promising results show that fast spectroscopic identification of single virus particles with TERS is possible.

Biosensors are able to look directly into the cell compartments—which used to be impossible without complex techniques—without affecting cell viability. However, uncoated nanoparticles suffer from lack of specificity. While features of the spectra could be assigned to some types of biomolecules, such a simple device fails to determine the presence of particular molecules. Rich spectra of biomolecules make it difficult to determine all the components accurately—profusion of the bands, as well as their overlapping, hinder clarity of the spectrum. In order to increase specificity, it is better to detect molecules in an indirect way. In this case, SERS signal is collected from a small reporter molecule adsorbed on a nanoparticle. Molecular dyes are frequently used as reporter molecules; as a result, even higher intensities thanks to SERRS effect are obtained.

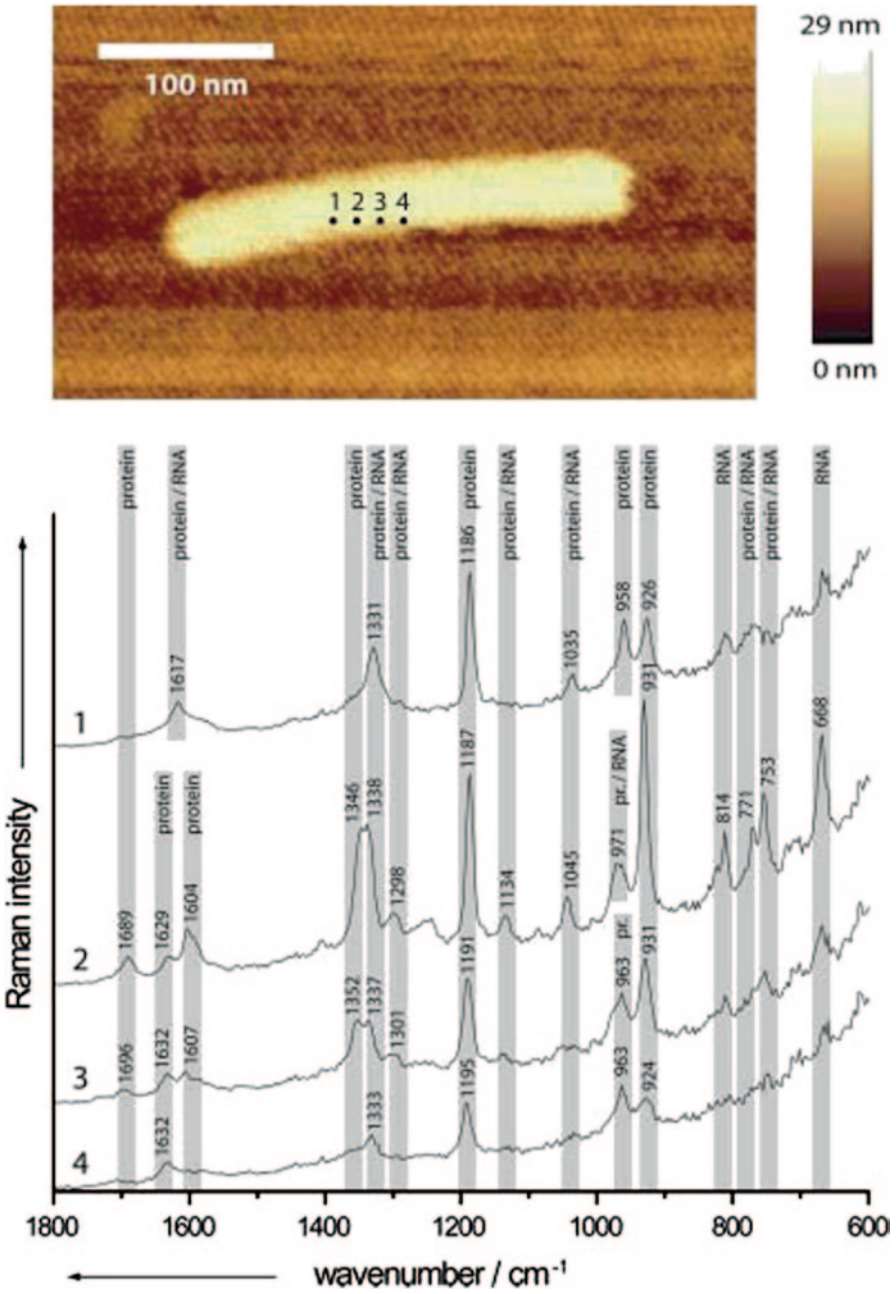


Fig. 3.9 *Top*: AFM image of tobacco mosaic virus (TMV); *bottom*: TERS spectra collected from four different points marked in the *top* picture. (Reproduced from Ref. [159] with kind permission of John Wiley and Sons)

Finding out the composition of the analyte is not always the main goal of the research. Sometimes it is more interesting to determine intracellular variables. Concept of the biosensor remains the same: surface of the nanoparticle yielding Raman enhancement is modified with reporter molecules whose SERS spectra are recorded. So that the sensor could monitor the environment, spectrum of the reporter should change with the studied parameter, such as intracellular redox potential. It plays an important role in protein interactions, apoptosis and signalling. In this approach, SERS spectra of the nanoprobe-attached molecules that exhibit SERS spectrum sensitive to oxidation state are recorded. Auchinvole et al. constructed a biosensor based on the SERS effect of two separate molecules, quinone derivatives: 1,8-diaza-4,5-dithia-1,8-di(2-chloro-[1,4]-naphthoquinone-3-yl)octane (NQ) and 2-mercaptobenzene-1,4-diol (HQ), deposited on the surface of gold nanoshells [162]. They undergo reversible two-electron redox reactions. The nanoprobes were introduced into single fibroblast cell and the SERS spectra were collected from the cytoplasm, where nanoprobes have been located. Plots of the marker band intensity changes show sigmoidal shape for both of the sensors. Based on the intensities of the respective SERS bands, relative content of the reduced/oxidized form of probe molecules was estimated, enabling determination of the value of resting cellular potential. What is important, it was proved that the introduction of the device into the cell does not change the redox potential inside. In separate experiments, the authors also demonstrated the capability of SERS to monitor the localized response to reductive and oxidative stress.

Another essential variable of the intracellular environment is the pH value. Different compartments are characterized by different pH ranges which additionally change over time. Anomalies in the pH value might be associated with several diseases, such as kidney failure or certain cancers. Importance of acidity inside the cell led to a big collection of pH biosensors throughout the literature. Determination of pH value inside the living cell may be an interesting illustration of possible applications of labeled SERS nanoprobes [163–170]. In these experiments, gold nanoparticles are functionalized with reporter molecules that demonstrate strong SERS spectrum sensitive to pH, such as 4-mercaptobenzoic acid (PMBA), 4-mercaptopyridine [165] or 4-aminothiophenol [167]. Next, the nanosensors are introduced into the cells and the SERS spectra of a reporter molecule are collected at many points of the sample, thus probing and imaging pH values inside the individual living cells. Due to changes in their chemical structure in varying pH, intensities of the chosen marker band vary as well. Kneipp et al. employed PMBA based sensor to perform SERS mapping of the pH-value in the cell [163]. They also broadened the working range of the sensor by exploiting SEHRS (surface-enhanced hyper-Raman scattering). A mobile SERS nanosensor made of aggregates of Au nanoparticles, functionalized with PMBA enabled monitoring changes in local pH value over time and observing some metabolic changes in living NIH/3T3 cells [166]. Local pH value was also monitored inside individual cancer cells after treatment by the photodynamic therapy drug emodine [168]. In this report, a micrometer-sized silica beads covered with Ag nanoparticles functionalized with PMBA were passively embedded into the cells. The pH changes over time revealed ability of the cell to

recover when emodin was activated with low light dosage (pH fluctuated slightly around the average value, about 5.5), while after treatment with a high light dosage, the intracellular pH increased substantially, triggering consecutive processes and finally cell death was observed.

This short review does not cover all applications of SERS in cellular sensing. For example, SERS nanotags were successfully applied to detect cancer markers in living cells [171, 172]. These reports clearly demonstrate perspectives of SERS spectroscopy as a promising tool in biomedicine.

3.5 Concluding Remarks

As shown in this review, SERS is an exceptionally selective and nondestructive spectroscopic method which offers sensitivity down to the single molecule level. It has a great potential for a variety of applications in bioscience, including *in situ* experiments in living cells and microorganisms. Despite great progress in bioanalytical applications of SERS spectroscopy in the last several years, there are still many obstacles that must be overcome to make this method a more universal tool. First, developing stable, reproducible, and highly enhancing substrates is still a great challenge in SERS spectroscopy of biological samples. Controlling the distribution of the SERS nanoparticles inside the living biological materials such as cells, tissues and microorganisms is another task that must be solved. In targeted sensing in biological samples, the magnetic plasmonic nanoparticles have great potential, since they enable displacing the nanoprobes in magnetic field. Developing biocompatible and non-toxic SERS nanoprobes is a very important and still unresolved problem in applications of SERS for *in vivo* experiments with living cells and microorganisms as well.

In the last decade, great progress has been made in combining Raman spectroscopy with scanning tunnelling (STM) or atomic force microscopy (AFM). It resulted in developing fascinating technique called TERS. This technique opens new, exciting perspectives for noninvasive imaging of a variety of biological systems, e.g. cell membranes with nanometric spatial resolution. Main challenge of TERS spectroscopy is fabrication of reproducible, high quality tips, that are necessary for effective enhancement of the Raman signal.

References

1. Fleischmann M, Hendra PJ, McQuillan AJ (1974) Raman spectra of pyridine adsorbed at a silver electrode. Chem Phys Lett 26(2):163–166
2. Jeanmaire DL, Van Duyne RP (1977) Surface Raman spectroelectrochemistry. Part I. Heterocyclic, aromatic, and aliphatic amines adsorbed on the anodized silver electrode. J Electroanal Chem Interfacial Electrochem 84(1):1–20
3. Lombardi JR, Birke RL, Lu T, Xu J (1986) Charge-transfer theory of surface enhanced Raman spectroscopy: Herzberg-Teller contributions. J Chem Phys 84:4174–4180
4. Pines D (1956) Collective energy losses in solids. Rev Mod Phys 28:184–198

5. Stern EA, Ferrell RA (1960) Surface plasma oscillations of a degenerate electron gas. Phys Rev 120:130–136
6. Le Ru EC, Blackie E, Meyer M, Etchegoin PG (2007) Surface enhanced Raman scattering enhancement factors: a comprehensive study. J Phys Chem C 111:13794–13803
7. Kottmann JP, Martin OJF, Smith DR, Schultz S (2001) Dramatic localized electromagnetic enhancement in plasmon resonant nanowires. Chem Phys Lett 341:1–6
8. Creighton JA (1988) The selection rules for surface-enhanced Raman spectroscopy. In: Clark RJH, Hester RE (eds) Spectroscopy of surfaces. Advances in Spectroscopy, vol 16. Wiley, New York, pp 37–89
9. Creighton A, Blatchford CG, Albrecht MG (1979) Plasma resonance enhancement of Raman scattering by pyridine adsorbed on silver and gold sol particles of size comparable to the excitation wavelength. J Chem Soc Faraday Trans 2 75:790–798
10. Lee PC, Meisel D (1982) Adsorption and surface-enhanced Raman of dyes on silver and gold sols. J Phys Chem 86:3391–3395
11. Leopold N, Lendl B (2003) A new method for fast preparation of highly surface-enhanced Raman scattering (SERS) active silver colloids at room temperature by reduction of silver nitrate with hydroxylamine hydrochloride. J Phys Chem B 107(24):5723–5727
12. Blatchford CG, Campbell JR, Creighton JA (1982) Plasma resonance enhanced Raman scattering by adsorbates on gold colloids: the effects of aggregation. Surf Sci 120:435–455
13. Larmour IA, Faulds K, Graham D (2012) SERS activity and stability of the most frequently used silver colloids. J Raman Spectrosc 43:202–206
14. Chumanov G, Sokolov K, Gregory BW, Cotton TM (1995) Colloidal metal films as a substrate for surface-enhanced spectroscopy. J Phys Chem 99:9466–9471
15. Grabar KC, Brown KR, Keating CD, Stranick SJ, Tang SL, Natan MJ (1997) Nanoscale characterization of gold colloid monolayers: a comparison of four techniques. Anal Chem 69:471–477
16. Fan M, Brolo AG (2009) Silver nanoparticles self assembly as SERS substrates with near single molecule detection limit. Phys Chem Chem Phys 11:7381–7389
17. Andrade GFS, Fan MK, Brolo AG (2010) Multilayer silver nanoparticles-modified optical fiber tip for high performance SERS remote sensing. Biosens Bioelectron 25:2270–2275
18. Addison CJ, Brolo AG (2006) Nanoparticle-containing structures as a substrate for surface-enhanced Raman scattering. Langmuir 22:8696–870
19. Li X, Xu W, Zhang J, Jia H, Yang B, Zhao B, Li B, Ozaki Y (2004) Self-assembled metal colloid films: two approaches for preparing new SERS active substrates. Langmuir 20:1298–1304
20. Tantra R, Brown RJC, Milton MJT, Gohil DD (2008) A practical method to fabricate gold substrates for surface-enhanced Raman spectroscopy. Appl Spectrosc Vol. 62:992–1000
21. Yang C, Xie YT, Yuen MMF, Xiong XM, Wong CP (2010) A facile chemical approach for preparing a SERS active silver substrate. Phys Chem Chem Phys 12(43):14459–14461
22. Semin DJ, Rowlen KL (1994) Influence of vapor deposition parameters on SERS active Ag film morphology and optical properties. Anal Chem 66:4324–4331
23. Roark SE, Rowlen KL (1994) Thin Ag films: influence of substrate and postdeposition treatment on morphology and optical properties. Anal Chem 66:261–270
24. Haynes CL, Van Duyne RP (2001) Nanosphere lithography: a versatile nanofabrication tool for studies of size-dependent nanoparticle optics. J Phys Chem B 105:5599–5611
25. Haes AJ, Chang L, Klein WL, Van Duyne RP (2005) Detection of a biomarker for Alzheimer's disease from synthetic and clinical samples using a nanoscale optical biosensor. J Am Chem Soc 127(7):2264–2271
26. Stosch R, Yaghobian F, Weimann T, Brown RJC, Milton MJT, Guttler B (2011) Lithographical gap-size engineered nanoarrays for surface-enhanced Raman probing of biomarkers. Nanotechnology 22(10):105303
27. Abu Hatab NA, Oran JM, Sepaniak MJ (2008) Surface-enhanced Raman spectroscopy substrates created via electron beam lithography and nanotransfer printing. ACS Nano 2(2):377–385

28. Cialla D, Huebner U, Schneidewind H, Moeller R, Popp J (2008) Probing innovative microfabricated substrates for their reproducible SERS activity. ChemPhysChem 9(5):758–762
29. Porter Jr LA, Choi HC, Schmeltzer JM, Ribbe AE, Elliott LCC, Buriak JM (2002) Electroless nanoparticle film deposition compatible with photolithography, microcontact printing, and dip-pen nanolithography patterning technologies. NanoLetters 2(12):1369–1372
30. Dieringer JA, McFarland AD, Shah NC, Stuart DA, Whitney AV, Yonzon CR, Young MA, Zhang XY, Van Duyne RP (2006) Surface enhanced Raman spectroscopy: new materials, concepts, characterization tools, and applications. Faraday Discussions 132:9–26
31. Baia L, Baia M, Popp J, Astilean S (2006) Gold films deposited over regular arrays of polystyrene nanospheres as highly effective SERS substrates from visible to NIR. J Phys Chem B 110(47):23982–23986
32. Haynes CL, Van Duyne RP (2003) Plasmon-sampled surface-enhanced Raman excitation spectroscopy. J Phys Chem B 107:7426–7433
33. Cialla D, März A, Böhme R, Theil F, Weber K, Schmitt M, Popp J (2012) Surface-enhanced Raman spectroscopy (SERS): progress and trends. Anal Bioanal Chem 403:27–54
34. Mohapatra S, Siddhanta S, Kumar DR, Narayana C, Maji TK (2010) Facile and green synthesis of SERS active and ferromagnetic silver nanorods. Eur J Inorg Chem 31:4969–4974
35. Smitha SL, GopchandranKG, Ravindran TR, Prasad VS (2011) Gold nanorods with finely tunable longitudinal surface plasmon resonance as SERS substrates. Nanotechnology 22:265705
36. Rycenga M, Camargo PHC, Li W, Moran CH, Xia Y, (2010) Understanding the SERS effects of single silver nanoparticles and their dimers, one at a time. J Phys Chem Lett 1:696–703
37. Sun Y, Xia Y (2002) Shape-controlled synthesis of gold and silver nanoparticles. Science 298:2176–2179
38. Rycenga M, Kim MH, Camargo PHC, Cobley C, Li Z-Y, Xia Y (2009) Surface-enhanced Raman scattering: comparison of three different molecules on single-crystal nanocubes and nanospheres of silver. J Phys Chem A 113:3932–3939
39. Eguchi M, Mitsui D, Wu H-L, Sato R, Teranishi T (2012) Simple reductant concentration-dependent shape control of polyhedral gold nanoparticles and their plasmonic properties. Langmuir 28(24):9021–9026
40. Khoury CG, Vo-Dinh T (2008) Gold nanostars for surface-enhanced Raman scattering: synthesis, characterization and optimization. J PhysChem C 112(48):18849–18859
41. Esenturk EN, Walker ARH (2009) Surface-enhanced Raman scattering spectroscopy via gold nanostars. J Raman Spectrosc 40(1):86–91
42. Rodriguez-Lorenzo L, Alvarez-Puebla RA, Garcia J, de Abajo F, Liz-Marzan LM (2010) Surface enhanced Raman scattering using star-shaped gold colloidal nanoparticles. J Phys Chem C 114(16):7336–7340
43. Kim JH, Kang T, Yoo SM, Lee SY, Kim B, Choi YK (2009) A well-ordered flower-like gold nanostructure for integrated sensors via surface-enhanced Raman scattering. Nanotechnology 20(23):235302
44. Gellner M, Kustner B, Schlucker S (2009) Optical properties and SERS efficiency of tunable gold/silver nanoshells. Vib Spectrosc 50:43–47
45. Oldenberg SJ, Averitt RD, Westcott SL, Halas NJ (1998) Nanoengineering of Optical Resonances. Chem Phys Lett 288:243–247
46. Alvarez-Puebla RA, Ross DJ, Nazri GA, Aroca RF (2005) Surface-enhanced Raman scattering on nanoshells with tunable surface plasmon resonance. Langmuir 21:10504–10508
47. Talley CE, Jackson JB, Oubre C, Grady NK, Hollars CW, Lane SM, Huser TR (2005) Surface-enhanced Raman scattering from individual Au nanoparticles and nanoparticle dimer substrates. Nano Lett 5:1569–1574
48. Jackson JB, Halas NJ (2004) Surface-enhanced Raman scattering on tunable plasmonic nanoparticle substrates. Proc Natl Acad Sci USA 101:17930–17935
49. Ochsenkuhn MA, Jess PRT, Stoquert H, Dholakia K, Campbell CJ, (2009) Nanoshells for surface-enhanced Raman spectroscopy in eukaryotic cells: cellular response and sensor development. ACS Nano 3(11):3613–3621

50. Lal S, Grady NK, Goodrich GP, Halas NJ (2006) Profiling the near field of a plasmonic nanoparticle with Raman-based molecular rulers. Nano Lett 6:2338–2343
51. Gellner M, Küstner B, Schlücker S (2009) Optical properties and SERS efficiency of tunable gold/silver nanoshells. Vibr Spectr 50:43–47
52. Li JF, Huang YF, Ding Y, Yang ZL, Li SB, Zhou XS, Fan FR, Zhang W, Zhou ZY, Wu DY, Ren B, Wang ZL, Tian ZQ (2010) Shell-isolated nanoparticle-enhanced Raman spectroscopy. Nature 64(7287):392–395
53. Anema JR, Li JF, Yang ZL, Ren B, Tian ZQ (2011) Shell-isolated nanoparticle-enhanced Raman spectroscopy: expanding the versatility of surface-enhanced Raman scattering. Annu Rev Anal Chem 4:129–150
54. Kustner B, Gellner M, Schutz M, Schoppler F, Marx A, Strobel P, Adam P, Schmuck C, Schlucker S (2009) SERS labels for red laser excitation: silica-encapsulated SAMs on tunable gold/silver nanoshells. Angew Chem Int Ed 48:1950–1953
55. Mulvaney SP, Musick M D, Keating CD. Natan MJ (2003) Glass-coated, analyte-tagged nanoparticles: a new tagging system based on detection with surface-enhanced Raman scattering. Langmuir 19:4784–4790
56. Merican Z, Schiller TL, Hawker CJ, Fredericks PM, Blakey I (2007) Self-assembly and encoding of polymer-stabilized gold nanoparticles with surface-enhanced Raman reporter molecules. Langmuir 23:10539–10545
57. Moskovits M (2005) Surface-enhanced Raman spectroscopy: a brief retrospective. J Raman Spectrosc 36:485–496
58. Halas NJ, Lal S, Chang WS, Link S, Nordlander P (2011) Plasmons in strongly coupled metallic nanostructures. Chem Rev 111:3913–3939
59. Camden JP, Dieringer JA, Wang Y, Masiello DJ, Marks LD, Schatz GC, Van Duyne RP (2008) Probing the structure of single-molecule surface-enhanced Raman scattering hot spots. J Am Chem Soc 130:12616–12617
60. Rycenga M, Camargo PHC, Li W, Moran CH, Xia Y (2010) Understanding the SERS effects of single silver nanoparticles and their dimers, one at a time. J Phys Chem Lett 1:696–703
61. Wustholz KL, Henry AI, McMahon JM, Freeman RG, Valley N, Piotti ME, Natan MJ, Schatz GC, Van Duyne RP (2010) Structure-activity relationships in gold nanoparticle dimers and trimers for surface-enhanced Raman spectroscopy. J Am Chem Soc 132:10903–10910
62. Stockle RM, Suh YD, Deckert V, Zenobi R (2000) Nanoscale chemical analysis by tip-enhanced Raman spectroscopy. Chem Phys Lett 318:131–136
63. Pettinger B, Picardi G, Schuster R, Ertl G (2003) Surface-enhanced and STM tip-enhanced Raman spectroscopy of CN ions at gold surfaces. J Electroanal Chem 554:293–299
64. Picardi G, Nguyen Q, Schreiber J, Ossikovski R (2007) Comparative study of atomic force mode and tunneling mode tip-enhanced Raman spectroscopy. Eur Phys J Appl Phys 40:197–201
65. Zhang W, Yeo BS, Schmid T, Zenobi R (2007) Single-molecule tip-enhanced Raman spectroscopy with silver tips. J Phys Chem C 111:1733–1738
66. Sonntag MD, Klingsporn JM, Garibay LK, Roberts JM, Dieringer JA, Seideman T, Scheidt KA, Jensen L, Schatz GC, Van Duyne RP (2012) Single-molecule tip-enhanced Raman spectroscopy. J Phys Chem C 116:478–483
67. Yeo BS, Stadler J, Schmid T, Zenobi R, Zhang W (2009) Tip-enhanced Raman spectroscopy—its status, challenges and future directions. Chem Phys Lett 472:1–13
68. Bailo E, Deckert V (2008) Tip-enhanced Raman scattering. Chem Soc Rev 37:921–930
69. Deckert-Gaudig T, Deckert V (2010) Tip-enhanced Raman scattering (TERS) and high-resolution bio nano-analysis—a comparison. Phys Chem Chem Phys 12:12040–12049
70. Cotton TM (1988) The application of surface-enhanced Raman scattering to biochemical systems. In: Clark RJH, Hester RE (eds) Spectroscopy of surfaces. Advances in Spectroscopy, vol 16. Wiley, New York, pp 91–153
71. Aliaga AE, Osorio-Roman I, Leyton P, Garrido C, Carcamo J, Caniulef C, Celis F, Díaz GF, Clavijo E, Gomez-Jeria JS, Campos-Vallette MM (2009) Surface-enhanced Raman scattering study of L-tryptophan. J Raman Spectrosc 40:164–169

72. Aliaga AE, Osorio-Roman I, Garrido C, Leyton P, Carcamo J, Clavijo E, Gomez-Jeria JS, Diaz FG, Campos-Vallette MM (2009) Surface enhanced Raman scattering study of L-lysine. Vib Spectrosc 50(1):131–135
73. Yaffe NR, Blanch EW (2008) Effects and anomalies that can occur in SERS spectra of biological molecules when using a wide range of aggregating agents for hydroxylamine-reduced and citrate-reduced silver colloids. Vibr Spectrosc 48:196–201
74. Suh JS, Moskovits M (1986) Surface-enhanced Raman spectroscopy of amino acids and nucleotide bases adsorbed on silver. J Am Chem Soc 108:4711–4718
75. Lee HI, Suh SW, Kim MS (1988) Raman spectroscopy of L-tryptophan-containing peptides adsorbed on a silver surface. J Raman Spectrosc 19:491–495
76. Vidugiris GJA, Gudavicius AV, Razumas VJ, Kulys JJ (1989) Structure-potential dependence of adsorbed enzymes and amino acids revealed by the surface enhanced Raman effect. Eur Biophys J 17:19–23
77. Stewart S, Fredericks PM (1999) Surface-enhanced Raman spectroscopy of amino acids adsorbed on an electrochemically prepared silver surface. Spectrochim Acta Part A 55:1641–1660
78. Podstawka E, Ozaki Y, Proniewicz L (2004) Part I: surface-enhanced Raman spectroscopy investigation of amino acids and their homodipeptides adsorbed on colloidal silver. Appl Spectrosc 58:570–580
79. Podstawka E, Ozaki Y, Proniewicz L (2005) Part III: surface-enhanced Raman spectroscopy investigation of amino acids and their homodipeptides deposited onto colloidal Gold Surface. Appl Spectrosc 59:1516–1526
80. Guerrero AR, Aroca RF (2012) Surface-enhanced Raman scattering of hydroxyproline. J Raman Spectrosc 43:478–481
81. Brolo AG, Germain P, Hager G (2002) Investigation of the adsorption of L-Cysteine on a polycrystalline silver electrode by surface-enhanced Raman scattering (SERS) and Surface-Enhanced Second Harmonic Generation (SESHG). J Phys Chem B 106:5982–5987
82. Lee HI, Kim MS, Suh SW (1991) Raman spectroscopy of sulphur-containing amino acids and their derivatives adsorbed on silver. J Raman Spectrosc 22:91–96
83. Graff M, Bukowska J (2005) Adsorption of enantiomeric and racemic cysteine on a silver electrode—SERS sensitivity to chirality of adsorbed molecules. J Phys Chem B 109:9567–9574
84. Lee HI, Kim MS, Suh SW (1991) Raman spectroscopy of sulphur-containing amino acids and their derivatives adsorbed on silver. J Raman Spectrosc 22:91–96
85. Kühnle A, Linderoth TR, Hammer B, Besenbacher F (2002) Chiral recognition in dimerization of adsorbed cysteine observed by scanning tunneling microscopy. Nature 415:891–893
86. Graff M, Bukowska J (2010) Enantiomeric recognition of phenylalanine by self-assembled monolayers of cysteine: surface-enhanced Raman scattering evidence. Vibr Spectr 52:103–107
87. Graff M, Bukowska J (2011) Surface-enhanced Raman Scattering (SERS) spectroscopy of enantiomeric and racemic methionine on a silver electrode—evidence for chiral discrimination in interactions between adsorbed molecules. Chem Phys Lett 509:58–61
88. Deckert-Gaudig T, Deckert V (2009) Ultraflat transparent gold nanoplates-ideal substrates for tip-enhanced Raman scattering experiments. Small 5(4):432–436
89. Deckert-Gaudig T, Deckert V (2009) Tip-enhanced Raman scattering studies of histidine on novel silver substrates. J Raman Spectrosc 40:1446–1451
90. Stewart S, Fredericks PM (1999) Surface-enhanced Raman spectroscopy of peptides and proteins adsorbed on an electrochemically prepared silver surface. Spectrochim Acta Part A 55:1615–1640
91. Wei F, Zhang D, Halas NJ, Hartgerink JD (2008) Aromatic amino acids providing characteristic motifs in the Raman and SERS spectroscopy of peptides. J Phys Chem B 112:9158–9164
92. Herne TM, Ahern A, Garrell RL (1991) Surface-enhanced Raman spectroscopy of peptides: preferential N-terminal adsorption on colloidal silver. J Am Chem Soc 113:846–854

93. Podstawka E, Ozaki Y (2008) Bombesin modified 6–14 C-terminal fragments adsorption on silver surfaces: influence of a surface substrate. Biopolymers 89:941–950
94. Podstawka E, Niaura G, Proniewicz LM (2010) Potential-dependent studies on the interaction between phenylalanine-substituted bombesin fragments and roughened Ag, Au, and Cu electrode surfaces. J Phys Chem B 114:1010–1029
95. Podstawka-Proniewicz E, Kudelski A, Younkyoo Kim Y, Proniewicz L (2012) Adsorption of neurotensin-family peptides on SERS-active Ag substrate. J Raman Spectrosc 43:1196–1203
96. Podstawka-Proniewicz E, Andrzejak M, Kafarski P, Kim Y, Proniewicz LM (2011) Vibrational characterization of L-valine phosphonate dipeptides: FT-IR, FT-RS, and SERS spectroscopy studies and DFT calculations. J Raman Spectrosc 42(5):958–979
97. Małek K, Królikowska A and Bukowska J (2012) Comparative studies on IR, Raman, and Surface Enhanced Raman scattering spectroscopy of dipeptides containing ΔAla and ΔPhe. J Phys Chem B 116:1414–1425
98. Keating CD, Kovaleski KM, Natan MJ (1998) Protein: colloid conjugates for surface enhanced Raman scattering: stability and control of protein orientation. J Phys Chem B 102(45):9404–9413
99. Pavel I, McCarney E, Elkhaled A, Morrill A, Plaxco K, Moskovits M (2008) Label-free SERS detection of small proteins modified to act as bifunctional linkers. J Phys Chem C 112:4880–4883
100. Drachev VP, Nashine VC, Thoreson MD, Ben-Amotz D, Davisson VJ, Shalaev VM (2005) Adaptive silver films for detection of antibody-antigen binding. Langmuir 21:8368–8373
101. Aroca R (2006) Surface-enhanced vibrational spectroscopy. John Wiley, Chichester, pp 91–96
102. Deckert-Gaudig T, Bailo E, Deckert V (2009) Tip-enhanced Raman scattering (TERS) of oxidised glutathione on an ultraflat gold nanoplate. Phys Chem Chem Phys 11:7360–7362
103. Love JC, Estroff LA, Kriebel JK, Nuzzo RG, Whitesides GM (2005) Self-assembled monolayers of thiolates on metals as a form of nanotechnology. Chem Rev 105:1103–1169
104. Moyano DF, Rotello VM, (2011) Nano Meets Biology: Structure and Function at the Nanoparticle Interface. Langmuir 27:10376–10385
105. Królikowska A, Bukowska J (2007) Self-assembled monolayers of mercaptosuccinic acid monolayers on silver and gold surfaces designed for protein binding. Part II: vibrational spectroscopy studies on cytochrome c immobilization. J Raman Spectrosc 38:943–949
106. Fedurco M (2000) Redox reactions of heme-containing metalloproteins:dynamic effects of self-assembled monolayers on thermodynamics and kinetics of cytochrome c electron-transfer reactions. Coord Chem Rev 209:263–331
107. Hobara D, Imabayashi Si, Kakiuchi T (2002) Preferential adsorption of horse heart cytochrome c on nanometer-scale domains of a phase-separated binary self-assembled monolayer of 3-mercaptopropionic acid and 1-hexadecanethiol on Au(111). Nano Lett 2:1021–1025
108. Dick LA, Haes AJ, Van Duyne RP (2000) Distance and orientation dependence of heterogeneous electron transfer: a surface-enhanced resonance Raman scattering study of cytochrome c bound to carboxylic acid terminated alkanethiols adsorbed on silver electrodes. J Phys Chem B 104:11752–11762
109. Bonifacio A, van der Sneppen L, Gooijer C, van der Zwan G (2004) Citrate-reduced silver hydrosol modified with *ö*-mercaptoalkanoic acids self-assembled monolayers as a substrate for surface-enhanced resonance Raman scattering. A study with cytochrome c. Langmuir 20:5858–5864
110. Królikowska A, Bukowska J (2010) Surface-enhanced resonance Raman spectroscopic characterization of cytochrome c immobilized on 2-mercaptoethanesulfonate monolayers on silver. J Raman Spectrosc 41:1621–1631
111. Murgida DH, Hildebrandt P (2004) Electron-transfer processes of cytochrome c at interfaces. New insights by surface-enhanced resonance Raman spectroscopy. Acc Chem Res 37:854–861

112. Grochol J, Dronov R, Lisdat F, Hildebrandt P, Murgida DH (2007) Electron transfer in SAM/Cytochrome/Polyelectrolyte hybrid systems on electrodes: a time-resolved surface-enhanced resonance Raman study. Langmuir 23:11289–11294
113. Hildebrandt P, Stockburger (1986) Surface-enhanced resonance Raman spectroscopy of cytochrome c at room and low temperatures. J Phys Chem 90:6017–6024
114. Oellerich S, Wackerbarth H, Hildebrandt P (2002) Spectroscopic characterization of non-native conformational states of cytochrome c. J Phys Chem B 106:6566–6580
115. Macdonald IDG, Smith WE (1996) Orientation of cytochrome c adsorbed on a citrate-reduced silver colloid surface. Langmuir 12:706–713
116. Murgida DH, Hildebrandt P (2001) Heterogenous electron transfer of cytochrome c on coated silver electrodes. Electric field effects on structure and redox potential. J Phys Chem B 105:1578–1586
117. Kranich A., Ly HK, Hildebrandt P, Murgida DH (2008) Direct observation of the gating step in protein electron transfer: electric-field-controlled protein dynamics. J Am Chem Soc 130:9844–9848
118. Yeo BS, Madler S, Schmid T, Zhang W, Zenobi R (2008) Tip-enhanced Raman spectroscopy can see more: the case of cytochrome c. J Phys Chem C 112:4867–4873
119. Larmour IA, Faulds K, Graham D (2010) The past, present and future of enzyme measurements using surface enhanced Raman spectroscopy. Chem Sci 1:151–160
120. Copeland RA, Fodor SPA, Spiro TG (1984) Surface-enhanced Raman spectra of an active flavo enzyme: glucose oxidase and riboflavin binding protein on silver particles. J Am Chem Soc 106:3872–3874
121. Broderick JB, Natan MJ, O'Halloran TV, Van Duyne RP (1993) Evidence for retention of biological activity of a non-heme iron enzyme adsorbed on a silver colloid: a surface-enhanced resonance Raman scattering study. Biochemistry 32:13771–13776
122. Wu ZS,. Zhou GZ, Jiang JH, Shen GL, Yu RQ (2006) Gold colloid-bienzyme conjugates for glucose detection utilizing surface-enhanced Raman scattering. Talanta 70:533–539
123. Dou X, Takama T, Yamaguchi Y, Yamamoto H, Ozaki Y (1997) Enzyme Immunoassay Utilizing Surface-Enhanced Raman Scattering of the Enzyme Reaction Product. Anal Chem 69:1492–1495
124. Stevenson R, Ingram A, Leung H, McMillan DC, Graham D (2009) Quantitative SERRS immunoassay for detection of human PSA. Analyst 134:842–844
125. Ruan CM, Wang W, Gu BH, (2006) Detection of alkaline phosphatase using surface-enhanced Raman spectroscopy. Anal Chem 78:3379–3384
126. Barhoumi A, Zhang D, Tam F., Halas NJ (2008) Surface-enhanced Raman spectroscopy of DNA. J Am Chem Soc 130:5523–5529
127. Shafer-Peltier KE, Haynes CL, Glucksberg MR, Van Duyne RP (2003) Towards a glucose biosensor based on surface-enhanced Raman scattering. J Am Chem Soc 125:588–593
128. Yonzon CR, Haynes CL, Zhang XY, Walsh JT, Van Duyne RP (2004) A glucose biosensor based on surface-enhanced Raman scattering: improved partition layer, temporal stability, reversibility, and resistance to serum protein interference. Anal Chem 76:78–85
129. Lyandres O, Shah NC, Yonzon CR, Walsh JT, Glucksberg MR, Van Duyne RP (2005) Real-time glucose sensing by surface-enhanced Raman spectroscopy in bovine plasma facilitated by a mixed decanethiol/mercaptohexanol partition layer. Anal Chem 77:6134–6139
130. Stuart DA, Yonzon CR, Zhang X, Lyandres O, Shah NC, Glucksberg MR, Walsh JT, Van Duyne RP (2005) Glucose sensing using near-infrared surface-enhanced Raman spectroscopy: gold surfaces, 10-day stability, and improved accuracy. Anal Chem 77:4013–4019
131. Zhang XY, Young MA, Lyandres O, Van Duyne RP (2005) Rapid detection of an anthrax biomarker by surface-enhanced Raman spectroscopy. J Am Chem Soc 127:4484–4489
132. Cheng H-W, Chen Y-Y, Lin X-X, Huan S-Y, Wu H-L, Shen G-L, Yu R-Q (2011) Surface-enhanced Raman spectroscopic detection of *Bacillus subtilis* spores using gold nanoparticle based substrates. Anal Chim Acta 707:155–163

133. Li M, Zhang J, Suri S, Sooter LJ, Ma D, Wu N (2012) Detection of adenosine triphosphate with an aptamer biosensor based on surface-enhanced Raman scattering. Anal Chem 84:2837–2842
134. Braun G, Lee SJ, Dante M, Nguyen T-Q, Moskovits M, Reich N (2007) Surface-enhanced Raman spectroscopy for DNA detection by nanoparticle assembly onto smooth metal films. J Am Chem Soc 129:6378–6379
135. Cao YC, Jin R, Mirkin CA (2002) Nanoparticles with Raman spectroscopic fingerprints for DNA and RNA detection. Science 297:1536–1540
136. Stokes RJ, Macaskill A, Dougan JA, Hargreaves PG, Stanford HM, Smith WE, Faulds K, Graham D (2007) Highly sensitive detection of dye-labelled DNA using nanostructured gold surfaces. Chem Commun 27:2811–2813
137. Faulds K, McKenzie F, Smith WE, Graham D (2007) Quantitative simultaneous multianalyte detection of DNA by dual-wavelength surface-enhanced resonance Raman scattering. Angew Chem Int Ed 46:1829–1831
138. Driskell JD, Kwarta KM, Lipert RJ, Porter MD (2005) Low-level detection of viral pathogens by a surface-enhanced Raman scattering based immunoassay. Anal Chem 77:6147–6154
139. Xu S, Ji X, Xu W, Li X, Wang L, Bai Y, Zhao B, Ozaki Y (2004) Immunoassay using probe-labelling immunogold nanoparticles with silver staining enhancement *via* surface-enhanced Raman scattering. Analyst 129:63–68
140. Yakes BJ, Lipert RJ, Bannantine JP, Porter MD (2008) Detection of mycobacterium avium subsp. paratuberculosis by a sonicate immunoassay based on surface-enhanced Raman scattering. Clin Vaccine Immunol 15:227–234
141. Wang G, Park HY, Lipert RJ (2009) Mixed monolayers on gold nanoparticle labels for multiplexed surface-enhanced Raman scattering based immunoassays. Anal Chem 81:9643–9650
142. Chourpa I, Lei FH, Dubois P, Manfaita M, Sockalingum GD (2008) Intracellular applications of analytical SERS spectroscopy and multispectral imaging. Chem Soc Rev 37:993–1000
143. Xie W, Su L, Shen A, Materny A, Hu J (2011) Application of surface-enhanced Raman scattering in cell analysis. J Raman Spectrosc 42:1248–1254
144. Kneipp J, Kneipp H, Kneipp K (2008) SERS—a single-molecule and nanoscale tool for bioanalytics. Chem Soc Rev 37:1052–1060
145. Vitol EA, Orynbayeva Z, Friedman G, Gogotsi Y (2012) Nanoprobes for intracellular and single cell surface-enhanced Raman spectroscopy (SERS). J Raman Spectrosc 43:817–827
146. Nabiev IR, Morjani H, Manfait M (1991) Selective analysis of antitumor drug interaction with living cancer cells as probed with surface-enhanced Raman spectroscopy Eur Biophys J 19:311–316
147. Kneipp J, Kneipp H, McLaughlin M, Brown D, Kneipp K (2006) *In Vivo* Molecular Probing of Cellular Compartments with Gold Nanoparticles and Nanoaggregates. Nano Letters. 6(10):2225–2231
148. Kneipp J, Kneipp H, Rajadurai A, Redmont RW, Kneipp K (2009) Optical probing and imaging of live cells using SERS labels. J Raman Spectrosc 40:1–5
149. Oyelere AK, Chen PC, Huang X, El-Sayed IH, El-Sayed MA (2007) Peptide-Conjugated Gold Nanorods for Nuclear Targeting. Bioconjugate Chem 18:1490–1497
150. Rodriguez-Lorenzo L, Krpetic Z, Barbosa S, Alvarez-Puebla RA, Liz-Marzan LM, Prior IA, Brust M (2011) Intracellular mapping with SERS-encoded gold stars. Integr Biol 3:922–926
151. Xie JP, Zhang QB, Lee JY (2008) The synthesis of SERS-active gold nanoflower tags for *In Vivo* applications. ACS Nano 2:2473–2480
152. Vitol EA, Orynbayeva Z, Bouchard MJ, Azizkhan-Clifford J, Friedman G, Gogotsi Y (2009) *In Situ* intracellular spectroscopy with surface enhanced Raman spectroscopy (SERS)-enabled nanopipettes. ACS Nano 3(11):3529–3536

153. Kneipp K, Haka AS, Kneipp H, Badizadegan K, Yoshizawa N, Boone C, Shafer-Peltier K, Motz JT, Dasari RR, Feld MS (2002) Surface-enhanced Raman spectroscopy in single living cells using gold nanoparticles. Appl Spectrosc 56:150–154
154. Jarvis RM, Law N, Shadi LT, O'Brien P, Lloyd JR, Goodacre R (2008) Surface-enhanced Raman scattering from intracellular and extracellular bacterial locations. Anal Chem 80:6741–6746
155. Efrima S, Zeiri L (2009) Understanding SERS of bacteria. J Raman Spectrosc 40:277–288
156. Gessner R, Winter C, Rosch P, Schmitt M, Petry R, Kiefer W, Lankers M, Popp J (2004) Identification of Biotic and Abiotic Particles by Using a Combination of Optical Tweezers and *In Situ* Raman Spectroscopy. ChemPhysChem 5:1159–1170
157. Scaffidi JP, Gregas MK, Seewaldt V, Vo-Dinh T (2009) SERS-based plasmonic nanobiosensing in single living cells. Anal Bioanal Chem 393:1135–1141
158. Bohme R, Richter M, Cialla D, Rosch P, Deckert V, Popp J (2009) Towards a specific characterization of components on a cell surface—combined TERS-investigations of lipids and human cells. J Raman Spectrosc 40:1452–1457
159. Cialla D, Deckert-Gaudig T, Budich C, Laue M, Moller R, Naumann D, Deckert V, Popp J (2009) Raman to the limit: tip-enhanced Raman spectroscopic investigations of a single tobacco mosaic virus. J Raman Spectrosc 40:240–243
160. Budich C, Neugebauer U, Popp J, Deckert V (2008) Cell Wall investigations utilizing tip-enhanced Raman scattering. J Microsc 229:533–539
161. Kozitskaya S, Deckert V, Schmitt M, Popp J (2007) Towards a detailed understanding of bacterial metabolism—spectroscopic characterization of *Staphylococcus Epidermidis*. ChemPhysChem 8:124–137
162. Auchinvole CAR, Richardson P, McGuinnes C, Mallikarjun V, Donaldson K, McNab H, Campbell CJ (2012) Monitoring intracellular redox potential changes using SERS nanosensors. ACS Nano 6(1):888–896
163. Kneipp J, Kneipp H, Wittig B, Kneipp K (2007) One- and two-photon excited optical pH probing for cells using surface-enhanced Raman and hyper-Raman nanosensors. Nano Lett 7:2819–2823
164. Talley CE, Jusinski L, Hollars CW, Lane SM, Huser T (2004) Intracellular pH sensors based on surface-enhanced Raman scattering. Anal Chem 76:7064–7068
165. Nowak-Lovato KL, Wilson BS, Rector KD (2010) SERS nanosensors that report pH of endocytic compartments during FcεRI transit. Anal Bioanal Chem 398:2019–2029
166. Kneipp J, Kneipp H, Wittig B, Kneipp K (2010) Following the dynamics of pH in endosomes of live cells with SERS nanosensors. J Phys Chem C 114:7421–7426
167. Zong S, Wang Z, Yang J, Cui Y (2011) Intracellular pH sensing using p-aminothiophenol functionalized gold nanorods with low toxicity. Anal Chem 83:4178–4183
168. Balint S, Rao S, Marro M, Miskovsky P, Petrov D (2011) Monitoring of local pH in photodynamic therapy-treated live cancer cells using surface-enhanced Raman scattering probes. J Raman Spectrosc 42:1215–1221
169. Bishnoi SW, Rozell CJ, Levin CS, Gheith MK, Johnson BR, Johnson DH, Halas NJ (2006) All-optical nanoscale pH meter. Nano Lett 6:1687–1692
170. Ando RA, Pieczonka NPW, Santos PS, Aroca RF (2009) Chromic materials for responsive surface-enhanced resonance Raman scattering systems: a nanometric pH sensor. Phys Chem Chem Phys 11:7505–7508
171. Qian X, Peng XH, Ansari DO, Yin-Goen Q, Chen GZ, Shin DM, Yang L, Young AN, Wang MD, Nie SM (2008) *In vivo* tumor targeting and spectroscopic detection with surface-enhanced Raman nanoparticle tags. Nat Biotech 26:83–90
172. Lee S, Kim S, Choo J, Shin SY, Lee YH, Choi HY, Ha S, Kang KH, Oh CH (2007) Biological imaging of HEK293 cells expressing PLCç1 using surface-enhanced Raman microscopy. Anal Chem 79:916–922

Chapter 4
Raman Optical Activity of Biological Samples

Katarzyna Chruszcz-Lipska and Ewan W. Blanch

Abstract In this Chapter, the fundamentals and instrumentation of vibrational Raman optical activity are briefly presented. Next, we describe selected examples of successful applications of computational approaches to the analysis of ROA spectra of biomolecules showing that such calculations are an important aspect of this incisive structural technique.

Keywords Chirality • Raman optical activity (ROA) • Theoretical calculations • Two-dimentional correlation analysis • Data clustering techniques • Terpenes • Carbohydrates

4.1 Introduction to the Theory of Raman Optical Activity

The fundamental scattering mechanism responsible for ROA was discovered in 1969 by Peter W. Atkins and Laurence D. Barron [1]. These researchers found that interference between light waves scattered via the molecular polarizability and optical activity tensors of a molecule leads to a dependence of the scattered light intensity on the degree of circular polarization of the incident light and also to a circular component in the scattered light. Barron along with A.D. Buckingham [1–12] subsequently published a more definitive version of the original theory in which they also introduced as an experimental observable a quantity they called the dimensionless circular intensity difference (CID), which was defined as:

$$\Delta = (I^{R} - I^{L}) / (I^{R} + I^{L}) \tag{4.1}$$

K. Chruszcz-Lipska (✉)
Jagiellonian Centre for Experimantal Therapeutics, Jagiellonian University,
Bobrzyńskiego 14 Str.,
30-348, Kraków, Poland
e-mail: lipska@chemia.uj.edu.pl

E. W. Blanch
Manchester Institute of Biotechnology and Faculty of Life Sciences,
The University of Manchester, 131 Princess Street,
M1 7DN, Manchester, UK

M. Baranska (ed.), *Optical Spectroscopy and Computational Methods in Biology and Medicine,* Challenges and Advances in Computational Chemistry and Physics 14,
DOI 10.1007/978-94-007-7832-0_4, © Springer Science+Business Media Dordrecht 2014

where I^{R} and I^{L} are the scattered Raman intensities in right- and left-circularly polarized incident light, respectively. CID expressions for *forward* (0°) and *backward* (180°) scattering geometries from an isotropic sample for incident transparent wavelengths much larger than the molecular dimensions can be expressed in terms of the electric dipole–electric dipole molecular polarizability tensor, $\alpha_{\alpha\beta}$, and the electric dipole–magnetic dipole and electric dipole–electric quadrupole optical activity tensors, $G'_{\alpha\beta}$ and $A_{\alpha\beta\gamma}$, respectively [13, 14], as [13]:

$$\Delta(0^{\circ}) = \frac{4[45\alpha G' + \beta(G')^2 - \beta(A)^2]}{c[45\alpha^2 + 7\beta(\alpha)^2]}, \tag{4.2a}$$

$$\Delta(180^{\circ}) = \frac{24\beta(G')^2 + \frac{1}{3}\beta(A)^2}{c45\alpha^2 + 7\beta(\alpha)^2}, \tag{4.2b}$$

where the isotropic invariants of these quantities are defined as

$$\alpha = \frac{1}{3}\alpha_{\alpha\alpha}, \tag{4.3a}$$

$$G' = \frac{1}{3}G'_{\alpha\alpha}, \tag{4.3b}$$

while the anisotropic invariants are defined as

$$\beta(\alpha)^2 = \frac{1}{2}(3\alpha_{\alpha\beta}\alpha_{\alpha\beta} - \alpha_{\alpha\alpha}\alpha_{\beta\beta}), \tag{4.4a}$$

$$\beta(G')^2 = \frac{1}{2}(3\alpha_{\alpha\beta}G'_{\alpha\beta} - \alpha_{\alpha\alpha}G'_{\beta\beta}), \tag{4.4b}$$

$$\beta(A)^2 = \frac{1}{2}\omega\alpha_{\alpha\beta}\varepsilon_{\alpha\gamma\delta}A_{\gamma\delta\beta}. \tag{4.4c}$$

For the Cartesian tensor notation used above, a repeated Greek suffix denotes summation over the three orthogonal components, and $\varepsilon_{\alpha\beta\gamma}$ is the third-rank antisymmetric unit tensor. If we consider the case of a molecule composed entirely of idealized axially-symmetric bonds, where $\beta(G')^2=\beta(A)^2$ and $\alpha G'=0$ [13, 15], a simple bond polarizability theory can be used to show that ROA is generated exclusively by *anisotropic* scattering, with the CID expressions simplifying to [13]

$$\Delta(0^{\circ}) = 0, \tag{4.5a}$$

$$\Delta(180^{\circ}) = \frac{32\beta(G')^2}{c[45\alpha^2 + 7\beta(\alpha)^2]}. \tag{4.5b}$$

Within this bond polarizability approximation ROA intensity is, therefore, not spatially symmetric but is of maximum intensity in backscattering and is zero in forward scattering. This presents a clear contrast to the case of conventional Raman scattering intensities as they are the same in forward and backward directions. This finding is also the reason why a backscattering geometry is favoured in ROA experiments, particularly for studies of biomolecules in aqueous solution [16, 17]. Although other scattering geometries have been used, notably right-angle scattering [REFS], the reduced intensities obtained make them less favourable than backscattering.

The use of superscripts "R" and "L" in equation 1 refers explicitly to the circular polarization states of the incident lightwave, and until around 2004 the great majority of experimental ROA spectra were recorded using this incident circular polarization (ICP) arrangement. However, ROA is also manifest as a small circularly polarized component in the scattered beam using incident light with a fixed polarization state, including unpolarized [1, 13, 18–20]. This scattered circular polarization (SCP) arrangement is signified by use of subscripts on the intensity symbols in equation 1 to denote the circular polarization components of scattered light, and is the most widely practiced form of ROA today. In the far from resonance approximation, SCP ROA provides equivalent information and spectra to those measured using ICP ROA. Other ROA measurement strategies are also possible and we direct interested readers to a number of references [18, 21, 22]. It should also be borne in mind that these equations apply specifically to Rayleigh (elastic) scattering. In the case of Raman (inelastic) scattering the same basic CID expressions apply but the molecular property tensors are substituted by the corresponding vibrational Raman transition tensors between the initial and final vibrational states, n_v and m_v. Therefore, $\alpha_{\alpha\beta}$ etc. are replaced by $\langle m_v \mid \alpha_{\alpha\beta}(Q) \mid n_v \rangle$ etc., where $\alpha_{\alpha\beta}(Q)$ etc. are effective polarizability and optical activity operators that depend parametrically on normal vibrational coordinates Q so that, within the Placzek polarizability theory of the Raman effect [23], ROA intensity depends on products of the type $(\partial\alpha_{\alpha\beta} / \partial Q)_0$ $(\partial G'_{\alpha\beta} / \partial Q)_0$ and $(\partial\alpha_{\alpha\beta} / \partial Q)_0 \varepsilon_{\alpha\gamma\delta}$ $(\partial A_{\gamma\delta\beta} / \partial Q)_0$.

In biopolymers, the normal vibrational modes are often highly complex, as they contain contributions from local vibrational coordinates within both the backbone and side chains. ROA cuts through the complexity presented by these modes in vibrational spectra because the most intense ROA signals are associated with vibrational modes that sample the most chiral and rigid structural elements. These modes are usually located within the backbone and so typically generate ROA band patterns that are highly informative about backbone conformation. By contrast, the parent conventional Raman spectra of proteins are usually dominated by bands from the side chains, which can make identification of Raman structural marker bands difficult.

The time scale of Raman scattering events is very fast and of the order of a few femtoseconds. This is much shorter than the timescales of conformational fluctuations, so that both Raman and ROA spectra are a superposition of individual 'snap-

shot' spectra from all of the distinct conformations present at equilibrium. However, as the ROA observables depend on absolute chirality, the spectral contributions from enantiomeric structures adopted by a mobile biopolymer exploring the full range of accessible conformations available tend to cancel each other out. Therefore, ROA exhibits a strong sensitivity to the dynamics of biomolecular structure and more conformationally mobile structures display considerably reduced ROA intensities compared to rigid biomolecules. This sensitivity contrasts with that of conventional Raman band intensities which are 'blind' to chirality and so are generally additive and consequently less sensitive to conformational mobility. Although ultraviolet circular dichroism (UVCD) also demonstrates an enhanced sensitivity to the dynamics of chiral structures, the level of sensitivity of UVCD is lower than that for ROA due to the former's dependence on electronic transitions.

4.2 Instrumentation and Measurement

As discussed above, a backscattering geometry is the most favourable one for routine measurement of ROA spectra of biomolecules in aqueous solution. Backscattering ROA measurements can be made using either the ICP, SCP or other strategies mentioned above, with the instrumentation required being different in each case.

Up until around 2004–2005, most of the reported ROA spectra of biomolecules were measured on ICP backscattering instruments at the University of Glasgow in the Barron laboratory. A detailed description of the optical layout of typical Glasgow backscattering ICP ROA instruments can be found elsewhere [24]. In summary, a visible argon ion laser beam at 514.5 nm or a frequency-doubled Nd:YAG laser beam at 532 nm is weakly focused into the sample solution contained in a small rectangular fused quartz cell. The cone of backscattered light is reflected off a 45° mirror, which has a small central hole drilled to allow passage of the incident laser beam, through an edge filter to remove the Rayleigh line and into the collection optics of a single grating spectrograph, a customized Kaiser Holospec fast imaging spectrograph. The detector is a Peltier-effect cooled back-thinned charge coupled device (CCD) camera with a quantum efficiency of ~80% over the spectral range used and operating in multichannel mode to allow the full spectral range to be measured in a single acquisition. Spectral acquisition is synchronized with the electro-optic modulator used to switch the state of polarization of the incident laser beam between right and left circular at a suitable rate. Counter-rotating half-wave plates inserted in the incident laser beam after the electro-optic modulater average the azimuths of any residual linear contaminants in the right- and left-circular polarization states of the beam and suppress artefacts arising from linear birefringence. Spectra are displayed in analog-to-digital converter units as a function of the Stokes Raman wavenumber shift with respect to the exciting laser line. Typical laser power at the sample is ~700 mW and sample concentrations of proteins, carbohydrates and nucleic acids are ~30–100 mg/ml while those of intact viruses are ~5–30 mg/ml. Under these conditions ROA spectra over the range ~600–1700 cm^{-1} are typically obtained in ~5–24 hours for proteins and nucleic acids and ~1–4 days for intact

viruses. ROA data below ~600 cm^{-1} are often unreliable from highly scattering samples due to offsets associate with the intense Rayleigh wing.

Although ICP ROA instruments of the type described above have established the value of ROA and will continue to be useful, a completely new design of ROA instrument with significant advantages inherent to the SCP strategy has recently been developed by Hug [19, 25]. In particular, 'flicker noise' arising from dust particles and density fluctuations in the sample, laser power fluctuations, etc. are eliminated as the intensity difference measurements required to extract the circularly polarized components of the scattered beam are taken between two orthogonal components of the scattered light measured simultaneously over the same acquisition period. This strategy cancels out the flicker noise, resulting in greatly superior signal-to-noise characteristics. The basic design is illustrated in Fig. 4.3 which corresponds to Hug's original implementation of the SCP strategy [25]; some of the details differ in later versions, but the basic principle is the same. The incident laser beam, the initial linear polarization state of which is 'scrambled' by a fast rotation of its azimuth, is deflected by a very small right-angle prism into the sample cell. The cone of backscattered light is collimated onto a liquid crystal retarder set to convert right- and left-circular polarization states into linear polarization states with azimuths perpendicular and parallel, respectively, to the plane of the instrument. A beam-splitting cube then diverts the perpendicular component at 90° to the propagation direction of the parallel component, which passes through undiverted. In this way, the right- and left-circularly polarized components of the backscattered light are separated and collected into the ends of two fiber optic cables. Each fiber optic converts the cross section from circular at the input end to a linear configuration at the output end that matches the entrance slit of the spectrograph, thereby enabling separate Raman spectra for the right- and left-circularly polarized components of the scattered light to be dispersed simultaneously, one above the other, on the CCD. The ICP ROA spectrum then corresponds to a difference in the small circularly polarized components in the Raman bands for a chiral molecule. Small differences in the two detection channels are compensated by interconverting their function through the switching of the liquid crystal retarder from the $-\lambda/4$ to the $+\lambda/4$ state. A commercial instrument based on this new design that also incorporates a sophisticated artefact suppression protocol, based on a 'virtual enantiomers' approach which greatly facilitates the routine acquisition of reliable ROA spectra [26], is now available (the Chiral*RAMAN* spectrometer from BioTools, Inc.).

4.3 Applications of Computational Approach to Analysis of ROA Spectra of Biomolecules

The first ROA spectra were published 40 years ago [2] and from that time molecular structures of hundreds of biomolecules were investigated in their native environments by vibrational Raman optical activity. Carbohydrates that play an essential role in various biological processes are one of many excellent examples that show the power of the ROA technique. They give prominent ROA signals over a wide

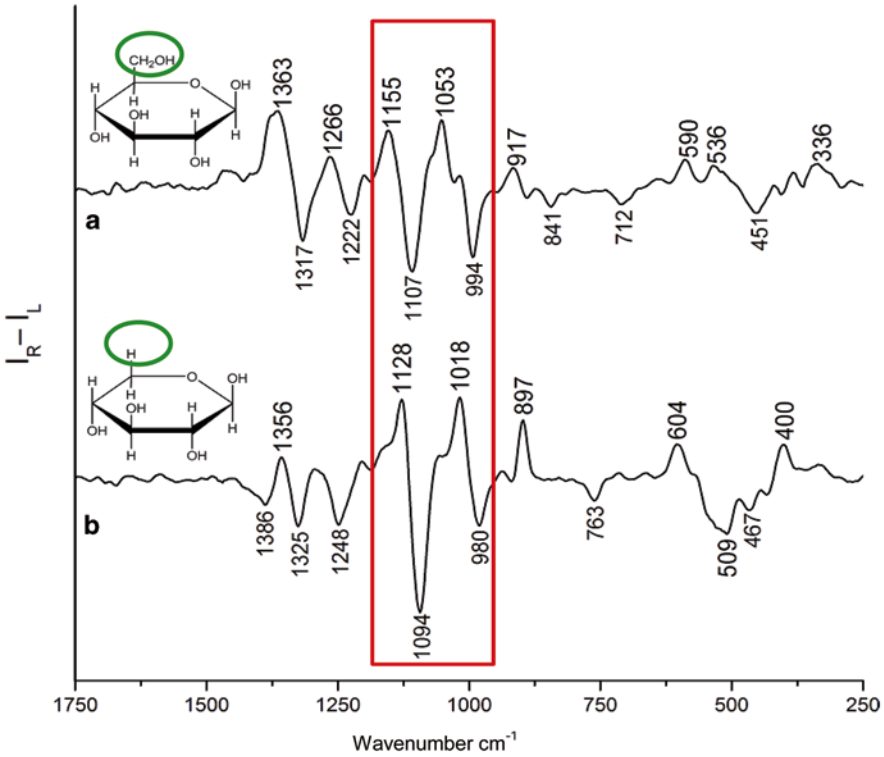


Fig. 4.1 ROA spectra of (**a**) D-glucose and (**b**) D-xylose illustrating the sensitivity of different spectral bands to the structural difference between these two monosaccharides, highlighted by the *green* rings. Both D-glucose and D-xylose occur in aqueous solution mainly as β-anomers

spectral range that are characteristic for the appropriate features of carbohydrate architecture. As was proposed in the literature [27, 28] the whole spectral range in the ROA spectra of monosaccharides can be divided into the regions that contain specific information about their stereochemistry: low frequency (250–600 cm^{-1}), anomeric (600–950 cm^{-1}), fingerprint (950–1200 cm^{-1}) and CH_2 and C–O–H deformations (1200–1500 cm^{-1}). Thanks to laborious analysis and detailed comparison of ROA spectra of monosaccharides from the same homomorphic series, including selected similar structures and isotopomers, it is possible to obtain information about conformation of the ring, the disposition of hydroxyl groups in the ring, the absolute configuration and vertical or horizontal orientation of groups bounded to the anomeric carbon, and as well the conformation of the exocyclic –CH_2OH group [27–29]. On the other hand, more complex compounds like di-, tri- [30–33], oligo-[34] or polysaccharides [35, 36] were successfully studied by ROA spectroscopy and gave additional information on glycosidic linkage types and its conformation and also secondary or even tertiary structure.

In Fig. 4.1 the ROA spectra of two homomorphic monosaccharides in aqueous solution: D-glucose and D-xylose are presented. D-glucose differs from D-xylose

only in that it has an exocyclic hydroxymethyl group. The rest of the two structures possess the same configuration at the individual chiral centres which is reflected in the similarity of the spectra in the range from 950–1200 cm^{-1}. As is seen in the fingerprint region of both spectra the characteristic sign pattern is observed, starting from high wavenumber for D-glucose: positive at 1155 cm^{-1}, negative at 1107 cm^{-1}, positive 1053 $cm^{-1}$and negative at 994 cm^{-1} and for D-xylose: positive at 1128 cm^{-1}, negative at 1094 cm^{-1}, positive at 1018 cm^{-1} and negative at 980 cm^{-1} [37]. In this region in the spectra of D-glucose in D_2O and also D-glucose-1-d the same sign pattern was found [27], although wavenumber shifts were observed in the spectrum of D-xylose. Figure 4.1. shows also that the spectra of these two monosaccharides differ significantly in other regions, which is due to the absence of the $-CH_2OH$ group in the structure of D-xylose.

Although the ROA spectra of carbohydrates are highly informative on their structure and many successful studies have been presented in the literature, the interpretation of these spectra is still difficult and limited, because there are no clear empirical rules relating the sign and magnitude of observed ROA signals to the molecular structure. Nowadays, we see that the most effective step towards complete understanding of ROA spectra is the use of quantum chemical calculations which are giving better results from year to year, starting from geometry prediction and ending with the computation of ROA observables [38–41]. In 2011, reliable calculations of the Raman and ROA spectra of methyl-β-D-glucose using a combination of density functional theory (DFT) and full molecular dynamics (MD) simulation of the aqueous environment were reported [42].

4.3.1 Calculation of ROA Spectra

Generally, the computational complexity of the molecular properties that are needed for simulation of an ROA spectrum and low experimental activity in ROA spectroscopy have been responsible for the rather slow development of methods for calculation of vibrational ROA spectra [43]. Calculations of vibrational Raman optical activity observables, usually based on the Placzek approximation, can be conducted in several ways. This multiplicity is a result of the existence of a variety of theoretical models of ROA [44–46]. Calculation by using such ROA models as the atom-dipole interaction [47] and the bond-polarizability [13, 48] didn't result in broad satisfactory results. However, they can still provide valuable physical insight into the generation of ROA [13, 45]. In 1990 the first complete *ab initio* calculation of a Raman optical activity spectrum was presented in the literature by Polavarapu and co-workers [49, 50] and in 2001 the first calculations of ROA CIDs at the DFT level of theory was made by Ruud et al. [51] which was a huge turning point in terms of presenting improvements in the quality of the obtained calculations. Nowadays, DFT theory has become the dominant approach for the calculation of vibrational Raman optical activity spectra.

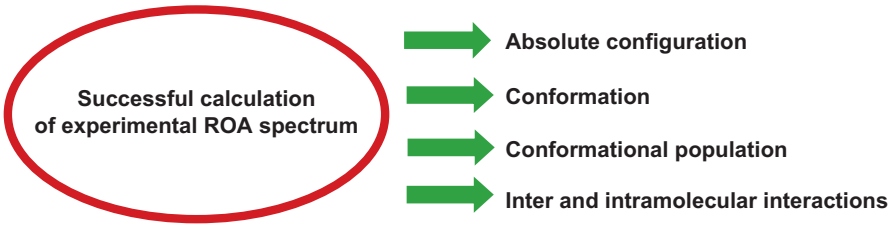


Fig. 4.2 Information obtainable from successful simulations of observed ROA spectra

Currently, ROA calculations may be performed using the Gaussian '09 [52] (Gaussian '03 [53] was the first commercially available user-friendly software for the DFT calculation of ROA spectra), Dalton [54], or Turbomole [55, 56] software packages and also programs like CADPAC [57] with which the first *ab initio* calculations of ROA were conducted. The calculation of the spectra using quantum-chemical methods is an important aspect of the ROA technique and is fast becoming the method of choice for obtaining detailed information about molecular structure from experimental ROA spectra (Fig. 4.2).

Successful simulation of observed spectra can afford the complete solution of the structure of the investigated compound, including its absolute configuration, conformation and, in the case of simultaneous existence of multiple conformers, also their percentage populations at the selected temperature. Additionally, these calculations give us the opportunity to obtain knowledge about inter- and intra-molecular interactions, especially between solvent molecules and the investigated compound.

Our purpose is not here to give a detailed and full account of the theoretical calculations, which can be found in other sources, including new review articles [43, 44, 58–61] and books [62], but rather to present some examples that show the possibility of application of DFT calculations in ROA spectroscopy of biological samples.

The calculation of vibrational ROA begins with building the structure that contains information specific to its stereochemical architecture. The next step is conducting the search over the conformational space to identify all of the lowest-energy conformers of the investigated molecule. This conformational search can be done in several ways. One of these is 'manual', the systematic specification of all angular degrees of freedom, including ring puckers and exploration of these degrees of freedom by an appropriate program. This approach is good for small systems, especially if solvent effects are not significant and conformational flexibility is limited like in the case of carenes, previously investigated by us (Scheme 1). Next, the set of lowest-energy conformers is usually optimized at a higher level of DFT theory using the basis set for the final ROA calculations. In the simplest case, only one conformer is predicted to exist, but in most cases a few or even many conformers significantly contribute to the equilibrium state. The conformers with an energy higher than the lowest-energy conformer by about 2 kcal (8.4 kJ) or less cover over 95% of the whole conformational population which is generally enough for the reliable calculation of vibrational spectra. Next, for each fully optimized conformer the vibrational frequencies and intensities are calculated and the total ROA spec-

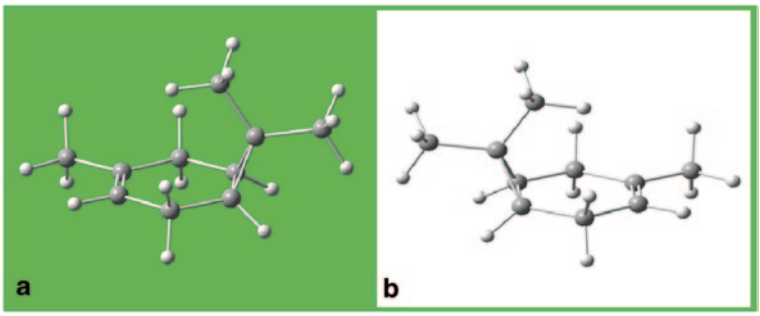
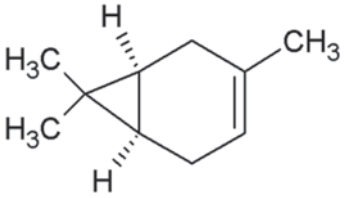


Fig. 4.3 Optimized geometries (DFT/B3LYP/aug-cc-pVTZ, Gaussian '09) of two enantiomers: (−)-3-carene (**a**) and (+)-3-carene (**b**)

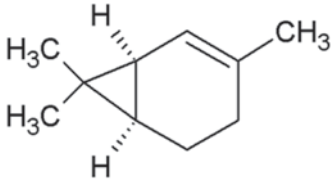
trum is produced as a weighted sum of component ROA spectra of the individual stable conformers, usually calculated according to the Boltzmann distribution. At this point, the simulated ROA spectrum can be compared with the measured one and analysed in a way appropriate to the application.

As a simple example of the calculation of ROA spectra we present our previous study of two bicyclic monoterpenes, 2- and 3-carene (Scheme 1), which occur naturally in plants. Delta-carene (3-carene), a major component of pine essential oil, has several biological activities such as anti-inflammatory [63], antibacterial [64], antifungal [65], inhibitory activity for acetylcholinesterase [66] and also anabolic activity in bone metabolism [67]. The experimental ROA spectra of 2- and 3-carene was reported in 1970s [68] but to the best of our knowledge ROA calculations had not been reported in the literature previously.

Scheme 1 Chemical structure of 2- and 3-carene



(+)-3-carene
(1S)-3,7,7-Trimethylbicyclo[4.1.0]hept-3-ene



(+)-2-carene
(1S)-3,7,7-Trimethylbicyclo[4.1.0]hept-2-ene

Exploration of the conformational space of (+)-3-carene by DFT calculations with the B3LYP functional and aug-cc-pVTZ basis set shows that only one conformer of that compound is stable in the gas phase at 298.15 K and its structure is presented in Fig. 4.3. Additionally, in Fig. 4.3 the optimized geometry of its enantiomer (−)-3-carene is presented.

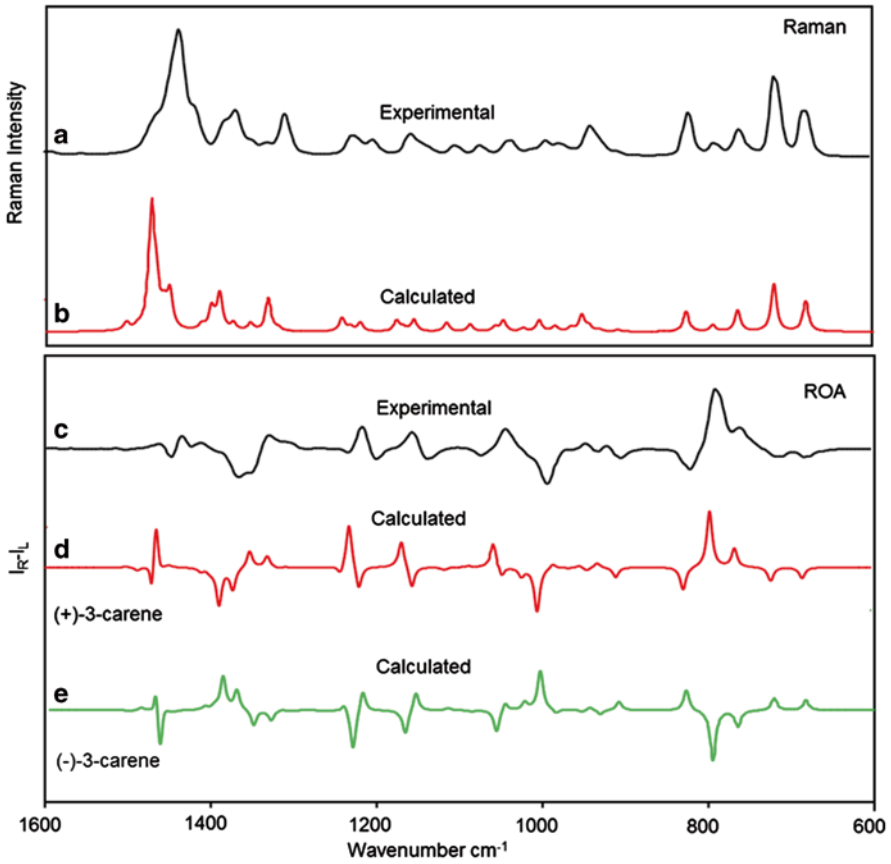


Fig. 4.4 Experimental and calculated spectra (DFT/B3LYP/aug-cc-pVTZ, Gaussian '09) of 3-carene. *Top*: experimental (neat *liquid*) (**a**) and calculated (**b**) Raman spectra, *bottom*: experimental (**c**) and calculated (**d**) Raman optical activity spectra of (+)-3-carene and additionally, calculated spectrum for (−)-3-carene (**e**). The spectral shapes were simulated using Lorentzian peaks with a bandwidth of 10 cm^{-1}. The calculated frequencies were not scaled

The comparison of experimental spectra of (+)-3-carene and calculated spectra for two enantiomers, i.e. (+)- and (−)-3-carene (Fig. 4.4), shows irrefutably that vibrational Raman optical activity is a powerful method for the determination of the absolute configuration of molecules. If the sign and relative magnitude of the observed bands in the ROA spectrum of the sample are the same as that of the simulated spectrum it means that the stereochemistry of the explored compound is the same as that chosen for the calculation. As we see in Fig. 4.4, the sign and relative intensity of observed bands in the measured and theoretical ROA spectra of (+)-3-carene are in close agreement. Thus, the geometry of (+)-3-carene is shown to be consistent with that presented in Fig. 4.3b.

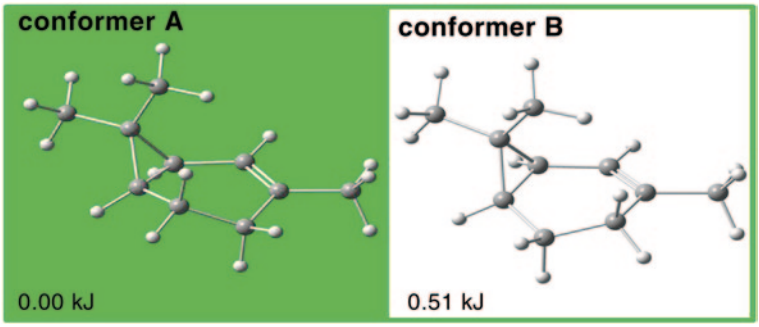


Fig. 4.5 Optimized geometries of most stable conformers of (+)-2-carene (DFT/B3LYP/aug-cc-pVTZ, Gaussian '09). The relative zero point corrected energy (in kJ mol^{-1}) at 298.15 K is given for each conformer

On the other hand, the comparison of calculated and observed spectra of 2-carene gives us information, not only about absolute configuration and conformation but also about the conformational population. The structures of the two lowest-energy conformers of (+)-2-carene differ in their ring conformations, which are shown along with their relative energies in Fig. 4.5.

According to our calculations, conformer A is slightly energetically favoured in the gas phase compared to conformer B.

In Fig. 4.6, a comparison between the experimental and calculated Raman optical activity spectra (DFT/B3LYP/aug-cc-pVTZ) of (+)-2-carene is presented. In this example we can appreciate the conformational sensitivity of Raman optical activity. As we see in Fig. 4.6, the calculated vibrational Raman optical activity spectra of conformers A and B of (+)-2-carene are completely different and it is difficult to find common features. Analysis of the experimental spectra of neat (+)-2-carene indicates that both conformers are present in the measured sample. In the experimental ROA spectrum of (+)-2-carene can be observed prominent bands at 1092 (negative), 1126 (positive) and 1252 cm^{-1} (negative) originating from conformer A, only. The equivalent theoretical bands are at about 1106 (negative), 1143 (positive) and 1264 cm^{-1} (negative), respectively.

On the other hand, very intensive bands only due to vibrations of conformer B are also observed in the experimental spectrum. The most significant signals appear at about 1303 (positive, cal. at 1322), 902 (positive, cal. at 908) and 700 cm^{-1} (negative, cal. at 704 cm^{-1}). Thus, the ROA experiment showed that two conformers of (+)-2-carene predicted by calculation in the gas phase in around a 1:2 ratio are also present in the neat liquid, but with equivalent quantity. As a result determination of the absolute configuration of (+)-2-carene is possible by comparison of experimental and calculated ROA spectra, with the latter being a sum of spectra of its two stable conformers (see Fig. 4.6).

In conclusion to this section we can state that nowadays it is possible to calculate routinely, and within a reasonable time, good quality ROA spectra for small and medium size molecules [69, 70].

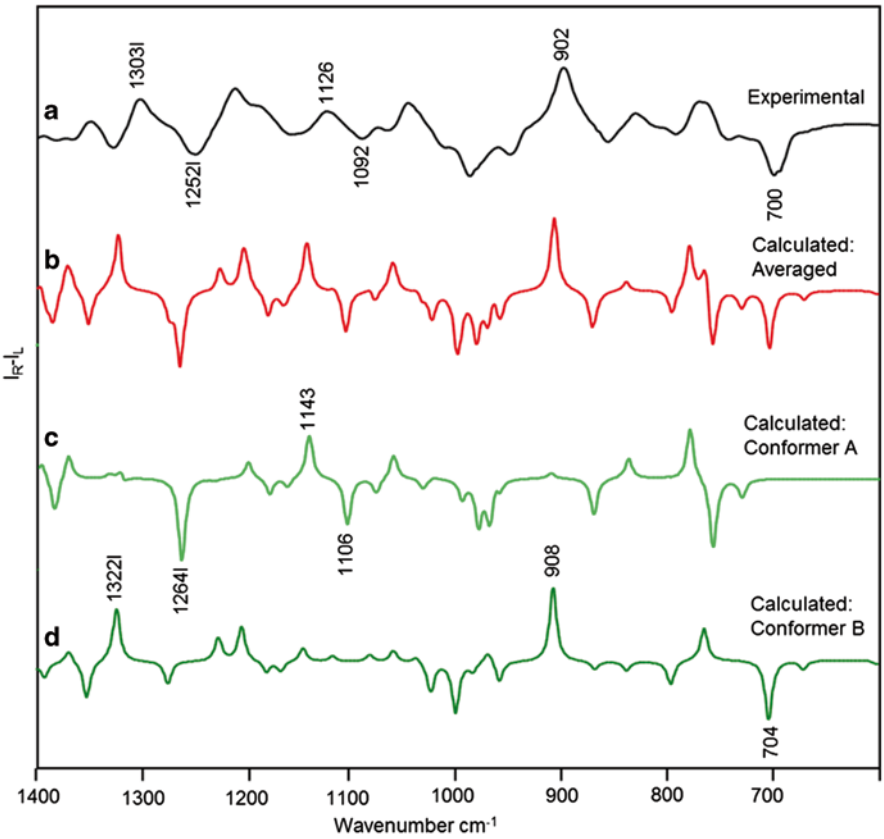


Fig. 4.6 Experimental and calculated Raman optical activity spectra (DFT/B3LYP/aug-cc-pVTZ, Gaussian '09) of (+)-2-carene: experimental (neat *liquid*) (**a**), calculated averaged (conformer A 58%, conformer B 42%) (**b**) calculated for conformer A (**c**) and for conformer B (**d**). The spectral shapes were simulated using Lorentzian peaks with a bandwidth of 10 cm^{-1}. The calculated frequencies ware not scaled

It is worth taking notice that the proper choice of the level of DFT theory is essential for the ROA calculation. Recent studies presented in the literature show that the basis set requirements for the calculation of the Raman/ROA tensor invariants and force field are different. Diffuse functions are important to compute properly the Raman/ROA tensors [69, 71–73] while for geometry optimization and force field calculations they usually are not required [69]. Prior analysis of the basis set dependence of vibrational Raman and ROA intensities has pointed out that calculations at the aug-cc-pVDZ or aug(sp)-cc-pVDZ levels of theory are the best option to obtain reasonable results for the 'one step' procedure in which the force field and Raman/ROA tensor invariant are computed at the same time using the same level of theory. For the more efficient, especially for large molecules, 'two step'

procedure in which the force field and Raman/ROA tensor invariants are computed in separate steps the following order is recommended: for small molecules aug-cc-pVDZ//cc-pVTZ (Raman/ROA tensor invariants//geometry and force fields), for medium size molecules aug(sp)-cc-pVDZ//cc-pVTZ or rDPS//6-31G*, and for large molecules the aug(sp)-cc-pVDZ//6-31G* or rDPS//6-31G* levels of theory [69]. It was also shown that ROA simulation carried out using the hybrid functional B3LYP [74] is a reasonable compromise balancing the quality of results and computer cost [75].

However, several challenges to the routine application of calculation of ROA spectra still remain. Conformational flexibility, size of investigated molecule and the strong effect of the solvent are the biggest problems to overcome. The total ROA spectrum can be very dependent on the population of the different conformers. In conformationally flexible molecules ROA signals arising from different conformations can cancel and so are not seen in the average ROA spectrum. Thus, correct determination of the relative energies for the relevant conformers is desirable for rendering realistic ROA spectra. However this is difficult to achieve, especially for molecules in solution when the influence of the solvent on the molecule is significant and needs to be taken into account. It is well known that a great advantage of the ROA method is the possibility of exploring the conformational dynamics of biomolecules in their native aqueous solution state. On the other hand, water is a capricious solvent to model, mainly because of its ability to make hydrogen bonds. In the current literature are presented several approaches to the modelling of solvent effects. Generally, they can be divided into implicit solvent models, where the solvent is defined as continuous medium and explicit solvent models, where the calculations are performed for a cluster containing the studied molecule surrounded by solvent molecules [60, 61, 76–80]. It has been recently shown that the use of the latter models can lead to better results for hydrated hydrogen bond-forming molecules [42, 76, 77].

Another challenging task is the calculation of ROA spectra for very large systems. In recent years full calculations for such big molecules have become possible and some scientists have taken up this challenge. So far, the largest systems interpreted by quantum-chemical procedures are helical decaalanine [81], valinomycin [82, 83], the β domain of metallothionein [84], a polypeptide of 20 alanine residues $(Ala)_{20}$ in an α-helical conformation [85] and the hormone insulin [86]. Until now, insulin is the biggest molecule for which an ROA spectrum was computed by quantum chemical calculations and this simulation predicted correctly the experimental spectroscopic response. The calculations on the insulin monomer and dimer were possible by using a popular approximation, the so-called Cartesian–coordinate tensor transfer method [87]. In this method the original large target molecule is divided into smaller molecular fragments for which the property tensors (force field, ROA polarizabilities, etc.) are calculated and in the next step are transferred back to the original molecule.

4.3.2 Multivariate Analysis of ROA Spectra

4.3.2.1 Two-dimensional Correlation Analysis

The ROA spectra of even simple biomolecules are complex and rich in detail. Although quantum chemical calculations are now opening many new insights into the structural origins of ROA features, these are not routine and are generally limited to small-to-medium sized molecules. In order to utilize the sensitivity of ROA to biomolecular structure to investigate conformational transitions, the informatics method of generalized two dimensional correlation analysis (2DCos) has proven to be a particularly promising tool. 2DCos provides a cross-correlation analysis of the measured spectra as a function of any two independent wavenumber positions, which spreads the spectra over an additional dimension. The generalized 2DCos method developed by Noda [88, 89] has been manifested in a wide, and still growing, range of representations, but the most common are 2D synchronous and asynchronous contour maps which show visual similarities to 2D NMR correlation maps.

Upon collection of a series of spectra from a sample subjected to varying degrees of a particular perturbation, 2D correlation generates a synchronous plot, in which relative similarities in behaviour exist between bands, and an asynchronous plot, where overall differences between the behaviour of bands are shown. Using simple sign associations, 2D spectral correlation analysis can provide much improved visualization of spectral features, consequently simplifying the identification of dynamic spectral changes, as well as probing the sequential information contained within the experimental data [90, 91]. 2D correlation also allows perturbation-induced data sets from two different spectroscopic techniques, such as corresponding Raman and ROA spectra measured simultaneously on the same sample, to be compared through so-called heterospectral or heterocorrelation spectroscopy. This can provide a source of new band assignments for ROA spectroscopists as the Raman literature is still more extensive.

Ashton et al. presented the first report of 2DCos ROA, in a study of the transition from α-helical to β-sheet structure in the homopolypeptide, poly-L-lysine [92]. A well known aspect of 2DCos, being both a key advantage and a potential source of error, is the enhanced sensitivity to small spectral changes, which necessitates the exercise of care in data pretreatment prior to performance of the 2D correlation analysis. Ashton et al. established an optimal protocol for pretreating ROA spectra in order to generate reliable synchronous and asynchronous correlation maps [92, 93]. The authors also found that the pseudoscalar nature of ROA spectra leads to an additional, but solvable, complication in the analysis of the contour plots. The general rules for interpreting the signs of 2D contour plots were based on monosignate spectra, e.g. Raman spectra, but the bisignate nature of ROA spectra needs to be taken into account in order to correctly identify the direction of intensity change. Recently, Pazderka and Kopecky have combined 2DCos with principal components analysis (PCA) in order to reduce noise and baseline distortions prior to generation of the 2D contour maps [94].

The ability of 2DCos ROA to reveal new structural information was shown by Ashton et al. in their study on α-helix unfolding in the model homopolypeptide poly-L-glutamic acid [95], which distinguished between fraying of the ends of helices and unfolding of the core helical structure.

4.3.2.2 Data Clustering Techniques

The analysis of protein structure remains a great challenge for the life sciences in the post-genomic era. Consequently, a large number of ROA experimental researches is dedicated to proteins and their building blocks, amino acids and peptides. Protein ROA spectra are dominated by bands from the peptide backbone and can give direct information about secondary and tertiary structure, whereas the bands from side chains are usually weaker because of some degree of conformational freedom. As a result, a lot of individual ROA signals and characteristic band patterns that are seen in protein spectra have been assigned to elements of secondary and tertiary structure such as α-helix and β-sheet, loops or turns. Thus, the ROA spectra of proteins with a large number of structure-sensitive bands are suitable for the application of pattern recognition methods to obtain useful information about the structure. Valuable structural information has been successfully found by studying protein ROA spectra using the method of principal component analysis (PCA) [96–100]. Thanks to this method the structural information from protein ROA spectra of unknown structure can be extracted automatically by their location into clusters corresponding to different protein folds. More advanced multivariate analysis using non-linear mapping (NLM) was found to give even better results [31, 101, 102]. A two-dimensional NLM plot for a set of 80 ROA spectra (both ICP and SCP) of polypeptides, proteins and viruses in aqueous solution measured in the range of 702–1773 cm^{-1} shows excellent clustering corresponding to the following seven structural categories: all α, all β, mainly α, mainly β, αβ, mainly disordered/irregular and all disordered/irregular [101]. Another two-dimensional NLM plot for a set of 85 ROA spectra revealed the significant differences between the structural characteristics of natively unfolded proteins and proteins unfolded by denaturation [102].

In conclusion, so far, applications of cluster analysis towards providing a better understanding of ROA spectra were used successfully for proteins, and showed that structural information can be easily and automatically extracted without any expert knowledge of characteristic band assignments. However, cluster analysis can be also used to study other class of biological molecules, as we have shown in our recent work where we used hierarchical cluster analysis to determine the content of chiral components in pichtae essential oil samples [103].

It is worth mentioning the application of another chemometric method, namely partial least-squares (PLS). These algorithms were also found to yield fine results for predictions of structural relationships among proteins from Raman as well as ROA spectra [104]. It was shown that the optimized PLS algorithm gives, for whole ROA spectra measured for 44 proteins, highly accurate secondary structure contents with correlation coefficients of 0.96–0.98 and RSMD values from 2.5 to 2.9 %. This

is considerably more accurate than has been found for analogous structural analyses using protein electronic circular dichroism (ECD) spectra, as ROA is a vibrational technique and so sensitive to far more structurally-sensitive bands.

4.4 Conclusions

No doubt, computational approaches to analysis of experimental ROA data can provide a great deal of information about the chiral structures of molecules. In the last decade, the considerable progress in quantum chemical calculation of ROA spectra has been achieved and the application of theoretical methods to better understand experimentally measured ROA spectra is becoming more widely used. Nowadays, not only are calculations of standard ROA spectra for various type of molecules well developed but there also now appear in the literature some examples of theoretical modelling of surface enhanced ROA (SEROA) [105] and resonance ROA (RROA) spectra [106, 107]. Both of these techniques have considerable potential for the investigation of biological samples because they combine chiral selectivity with surface or resonance enhancement of a measured signal. Although SEROA and RROA are in the early stage of their experimental development, the calculations reported follow these developments quite closely and already demonstrate the insight they can provide. It is also known in the literature that anharmonic effects can be important for ROA [108] but regardless of successes in that field, the theoretical calculation of ROA spectra can be still a challenging task.

Acknowledgements This research was partially supported by the British Council and the Polish Ministry of Science and Higher Education in 2010 (grant to K. Ch-L. from the British-Polish Young Scientist Programme). The authors thank the Academic Computer Centre 'Cyfronet' in Kraków (Poland) for computer time (grant no. KBN/SGI_ORIGIN_2000/UJ/061/2000). The authors also thank Zofia Barto for help in Fig. 4.1 preparation.

References

1. Atkins PW, Barron LD (1969) Rayleigh scattering of polarized photons by molecules. Mol Phys 16(5):453–466
2. Barron LD, Bogaard MP, Buckingham AD (1973) Raman scattering of circularly polarized light by optically active molecules. J Am Chem Soc 95:603–605
3. Barron LD, Hecht L, Blanch EW, Bell AF (2000) Solution structure and dynamics of biomolecules from Ramam optical activity. Prog Biophys Mol Biol 73:1–49
4. Barron LD, Blanch EW, Hecht L (2002) Unfolded proteins studied by Raman optical activity. Adv Protein Chem 62:51–90
5. Blanch EW, Hecht L, Barron LD (2003) Vibrational Raman optical activity of proteins, nucleic acids, and viruses. Methods 29:196–209
6. Hecht L, Phillips AL, Barron LD (1995) Determination of enantiomeric excess using vibrational Raman optical activity. J Raman Spectrosc 26:727–732

7. Spencer KM, Edmonds RB, Rauh RD, Carrabba MM (1994) Analytical determination of enantiomeric purity using Raman optical activity. Anal Chem 66:1269–1273
8. Spencer KM, Edmonds RB, Rauh RD (1996) Analytical chiral purity verification using Raman optical activity. Appl Spectrosc 50:681–685
9. Costante J, Hecht L, Polavarapu PL, Collet A, Barron LD (1997) Absolute-Configuration of bromochlorofluoromethane from experimental and Ab-Initio theoretical vibrational Raman optical-activity. Angew Chem Int Ed 36(8):885–887
10. Polavarapu PL (1998) Vibrational spectra: principles and applications with emphasis on optical activity. Elsevier, Amsterdam
11. Haesler J, Schindelholz L, Riguet E, Bochet CG, Hug W (2007) Absolute configuration of chirally deuterated neopentane. Nature 446:526–529
12. Barron LD, Buckingham AD (1971) Rayleigh and Raman scattering from optically active molecules. Mol Phys 20(6):1111–1119
13. Barron LD (2004) Molecular light scattering and optical activity, 2nd edn. Cambridge University Press, Cambridge
14. Barron LD, Buckingham AD (1974) Simple two-group model for Rayleigh and Raman optical activity. J Am Chem Soc 96:4769–4773
15. Buckingham AD (1967) Permanent and induced molecular moments and long-range intermolecular forces. Adv Chem Phys 12:107–142
16. Hecht L, Barron LD, Hug W (1989) Vibrational Raman optical activity in backscattering. Chem Phys Lett 158:341–344
17. Hecht L, Barron LD (1990) Analysis of modulation experiments for Raman optical activity. Appl Spectrosc 44:483–491
18. Nafie LA (1997) Infrared and Raman vibrational optical activity: theoretical and experimental aspects. Ann Rev Phys Chem 48:357–386
19. Hug W (2002) In: Chalmers JM, Griffiths PR (eds) Handbook of vibrational spectroscopy, Vol 1. Wiley, Chichester, p 745
20. Barron LD, Buckingham AD (1975) Rayleigh and Raman optical activity. Ann Rev Phys Chem 26:381–396
21. Nafie LA, Freedman TB (1989) Dual circular polarization Raman optical activity. Chem Phys Lett 154:260–266
22. Barron LD, Hecht L, McColl IH, Blanch EW (2004) Raman optical activity comes of age. Mol Phys 102:731–744
23. Long DA (2002) The Raman effect. Wiley, Chichester
24. Hecht L, Barron LD, Blanch EW, Bell AF, Day LA (1999) Raman optical activity instrument for studies of biopolymer structure and dynamics. J Raman Spectrosc 30:815–825
25. Hug W, Hangartner G (1999) A novel high-throughput Raman spectrometer for polarization difference measurements. J Raman Spectrosc 30:841–852
26. Hug W (2003) Virtual enantiomers as the solution of optical activity's deterministic offset problem. Appl Spectrosc 57:1–13
27. Bell AF, Barron LD, Heht L (1994) Vibrational Raman optical activity study of D-glucose. Carbohydr Res 257:11–24
28. Bell AF, Heht L, Barron LD (1995) Vibrational Raman optical activity of ketose monosaccharides. Spectrochim Acta A 51:1367–1378
29. Wen ZQ, Barron LD, Heht L (1993) Vibrational Raman optical activity of monosaccharides. J Am Chem Soc 115:285–292
30. Bell AF, Heht L, Barron LD (1994) Disaccharide solution stereochemistry from vibrational raman optical activity. J Am Chem Soc 116:5155–5161
31. Zhu AF, Isaacs NW, Hecht L, Tranter GE, Barron LD (2006) Raman optical activity of proteins, carbohydrates and glycoproteins. Chirality 18:103–115
32. Johannessen C, Pendrill R, Widmalm G, Hecht L, Barron LD (2011) Glycan structure of a high-Mannose glycoprotein from Raman optical activity. Angew Chem Int Ed 50:5349–5351
33. Barron LD, Gargaro AR, Wen ZQ (1991) Vibrational Raman optical activity of carbohydrates. Carbohydr Res 210:39–49

34. Bell AF, Heht L, Barron LD (1997) New evidence for conformational flexibility in cyclodextrins from vibrational Raman optical activity. Chem Eur J 3:1292–1298
35. Bell AF, Heht L, Barron LD (1995) Polysaccharide vibrational Raman optical activity: laminarin and pullulan. J Raman Spectrosc 26:1071–1074
36. Yaffe NR, Almond A, Blanch EW (2010) A new route to carbohydrate secondary and tertiary structure using Raman spectroscopy and Raman optical activity. J Am Chem Soc 132:10654–10655
37. Barto Z (2012) Badanie związków biologicznie czynnych metodą ramanowskiej aktywności optycznej (Study of biological active compounds by raman optical activity techniques). Dissertation, Jagiellonian University
38. Hoffman M, Rychlewski J (2001) Effects of substituting a OH group by a F atom in D-glucose. Ab initio and DFT analysis. J Am Chem Soc 123:2308–2316
39. Momany FA, Appell M, Strati G, Willett JL (2004) B3LYP/6-311 + + G** study of monohydrates of α- and β-d-glucopyranose: hydrogen bonding, stress energies, and effect of hydration on internal coordinates. Carbohydr Res 339:553–567
40. Corchado JC, Sanchez ML, Aguilar MA (2004) Theoretical study of the relative stability of rotational conformers of alpha and beta-D-glucopyranose in gas phase and aqueous solution. J Am Chem Soc 126:7311–7319
41. Macleod NA, Johannessen C, Hecht L, Barron LD, Simons JP (2006) From the gas phase to aqueous solution: Vibrational spectroscopy, Raman optical activity and conformational structure of carbohydrates. Int J Mass Spectr 253:193–200
42. Cheeseman JR, Shaik MS, Popelier PL, Blanch EW (2011) Calculation of Raman optical activity spectra of methyl-β-D-glucose incorporating a full molecular dynamics simulation of hydration effects. J Am Chem Soc 133:4991–4997
43. Ruud K, Thorvaldsen AJ (2009) Theoretical approaches to the calculation of Raman optical activity spectra. Chirality 21:E54–E67
44. Barron LD, Buckingham AD (2010) Vibrational optical activity. Chem Phys Lett 492:199–213
45. Barron LD, Vrbancich J (1984) Natural vibrational Raman optical activity. Springer, Berlin/Heidelberg
46. Hecht L, Barron LD (1996) Modern techniques in modern spectroscopy. Wiley
47. Prasad PL, Nafie LA (1979) The atom dipole interaction model of Raman optical activity: reformulation and comparison to the two-group model. J Chem Phys 70:5582–5588
48. Escribano JR, Barron LD (1988) Valence Optical Theory of Vibrational Circular Dichroism and Raman Optical Activity. Mol Phys 65(2):327–344
49. Polavarapu PL (1990) Ab Initio Raman and Raman optical activity spectra. J Phys Chem 94:8106–8112
50. Bose PB, Polavarapu PL, Hecht L (1990) Ab Initio and experimental Raman optical activity in (+)-®-methyloxirane. J Phys Chem 94:1734–1740
51. Ruud K, Helgaker T, Bour P (2002) Gauge-origin independent density-functional theory calculations of vibrational Raman optical activity. J Phys Chem A 106:7448–7455
52. Gaussian 09 Revision A.1 (2009) Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery Jr JA, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas Ö, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ Gaussian Inc Wallingford CT
53. Gaussian 03 Revision C02 (2004) Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery Jr JA, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA,

Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C and Pople JA, Gaussian Inc Wallingford CT

54. Dalton, a molecular electronic structure program, Release 1.2 (2001) Helgaker T, Jensen HJA, Jørgensen P, Olsen J, Ruud K, Ågren H, Auer AA, Bak KL, Bakken V, Christiansen O, Coriani S, Dahle P, Dalskov EK, Enevoldsen T, Fernandez B, Hättig C, Hald K, Halkier A, Heiberg H, Hettema H, Jonsson D, Kirpekar S, Kobayashi R, Koch H, Mikkelsen KV, Norman P, Packer MJ, Pedersen TB, Ruden TA, Sanchez A, Saue T, Sauer SPA, Schimmelpfennig B, Sylvester-Hvid KO, Taylor RP and Vahtras O
55. Ahlrichs R, Bar M, Haser M, Horn H, Kolmel C (1989) Electronic structure calculations on workstation computers: The program system Turbomole Chem Phys Lett 162(3):165–169
56. Luber S, Reiher M (2008) Raman optical activity spectra of chiral transition metal complexes Chem Phys 346:212–223
57. Amos RD, Alberts IL, Andrews JS, Colwell SM, Handy NC, Jayatilaka D, Knowles PJ, Kobayashi R, Laidig KE, Laming G, Lee AM, Maslen PE, Murray CW, Rice JE, Simandiras ED, Stone AJ, Su M-D, Tozer DJ (1995) CADPAC: the Cambridge analytic derivatives package issue 6 Cambridge
58. He Y, Wang B, Dukor R, Nafie LA (2011) Determination of absolute configuration of chiral molecules using vibrational optical activity: a review Appl Spectrosc 65:699–723
59. Barron LD (2012) From cosmic chirality to protein structure: Lord Kelvin's legacy Chirality 24:879–893
60. Pecul M (2009) New applications and challenges for computational ROA spectroscopy Chirality 21:E98–E104
61. Pecul M, Ruud K (2005) Ab initio calculation of vibrational Raman optical activity. Int J Quant Chem 104:816–829
62. Nafie LA (2011) Vibrational optical activity, principles and application. Wiley, Chichester
63. Ocete MA, Risco S, Zarzuelo A, Jimenez J (1989) Pharmacological activity of the essential oil of Bupleurum gibraltaricum: anti-inflammatory activity and effects on isolated rat uteri. J Ethnopharmacol 25:305–313
64. Pichette A, Laurouche PL, Lebrun M, Legault J (2006) Composition and antibacterial activity of Abies balsamea essential oil. Phytother Res 20:371–373
65. Cavaleiro C, Pinto E, Goncalves MJ, Salgueiro L (2006) Antifungal activity of Juniperus essential oils against dermatophyte, Aspergillus and Candida strains. J Appl Microbiol 100:1333–1338
66. Miyazawa M, Yamafuji C (2005) Inhibition of acetylcholinesterase activity by bicyclic monoterpenoids. J Agric Food Chem 53:1765–1768
67. Jeong J-G, Kim YS, Min YK, Kim SH (2008) Low concentration of 3-carene stimulates the differentiation of mouse osteoblastic MC3T3-E1 subclone 4 cells. Phytother Res 22:18–22
68. Barron LD, Clark BP (1979) Raman optical activity of pinenes, carenes, and related Molecules. JCS Perkin II:1171–1175
69. Cheeseman JR, Frisch MJ (2011) Basis set dependence of vibrational Raman and Raman optical activity intensities. J Chem Theory Comput 7:3323–3334
70. Haesler J, Schindelholz I, Riguet E, Bochet CG, Hug W (2007) Absolute configuration of chirally deuterated neopentane. Nature 446:526–529
71. Rappoport D, Furche FJ (2007) Lagrangian approach to molecular vibrational Raman intensities using time-dependent hybrid density functional theory. J Chem Phys 126:201104
72. Halls MD, Schlegel HB (1999) Comparison study of the prediction of Raman intensities using electronic structure methods. J Chem Phys 111:8819

73. Zuber G, Hug WJ (2004) Rarefied basis sets for the calculation of molecular optical tensors I The Importance of Gradients on Hydrogen Atoms for the Raman Scattering. Tensor Phys Chem A 108:2108–2118
74. Becke AD (1993) Density-functional thermochemistry III The role of exact exchange. J Chem Phys 98:5648–5652
75. Bour P (2001) Computations of the Raman optical activity via the Sum-Over-States expansions. J Comp Chem 22:426–435
76. Hopmann KH, Ruud K, Pecul M, Kudelski A, Dracinsky M, Bour P (2011) Explicit versus implicit solvent modeling of Raman optical activity Spectra. J Phys Chem B 115:4128–4137
77. Kaminski M, Kudelski A, Pecul M (2012) Vibrational Optical Activity of Cysteine in Aqueous Solution: A Comparison of Theoretical and Experimental Spectra. J Phys Chem B 116:4976–4990
78. Kapitan J, Baumruk V, Kopecky VJ, Bour P (2006) Conformational flexibility of L-alanine zwitterion determines shapes of Raman and Raman optical activity spectral bands. J Phys Chem A 110:4689–4696
79. Pecul M, Lamparska E, Cappelli C, Frediani L, Ruud K (2006) Solvent effects on Raman optical activity spectra calculated using the polarizable continuum model. J Phys Chem A 110:2807–2815
80. Hudecova J, Hornicek J, Budesinsky M, Sebestik J, Safarik M, Zhang G, Keiderling TA, Bour P (2012) Three types of induced tryptophan optical activity compared in model dipeptides: theory and experiment. Chem Phys Chem 13:2748–2760
81. Herrmann C, Ruud K, Reiher M (2006) Can Raman optical activity separate axial from local chirality? A theoretical study of helical deca-alanine. Chem Phys Chem 7:2189–2196
82. Yamamoto S, Watarai H, Bour P (2011) Monitoring the backbone conformation of Valinomycin by Raman Optical Activity. Chem Phys Chem 12:1509–1518
83. Yamamoto S, Straka M, Watarai H, Bour P (2010) Formation and structure of the potassium complex of valinomycin in solution studied by the Raman optical activity spectroscopy. Phys Chem Chem Phys 12:11021–11032
84. Luber S, Reiher M (2010) Theoretical Raman optical activity study of the β domain of rat metallothionein. J Phys Chem B 114:1057–1063
85. Jacob CR, Luber S, Reiher M (2009) Understanding the signatures of secondary-structure elements in proteins via Raman optical activity spectroscopy. Chem Eur J 15:13491–13508
86. Yamamoto B, Kaminsky J, Bour P (2012) Structure and vibrational motion of insulin from Raman optical activity spectra. Anal Chem 84:2440–2451
87. Bour P, Sopkova J, Bednarova L, Malon P, Keiderling TA (1997) Transfer of molecular property tensors in cartesian coordinates: a new algorithm for simulation of vibrational spectra. J Comput Chem 18:646–659
88. Noda I (1993) Generalized two-dimensional correlation method applicable to infrared, Raman, and other types of spectroscopy. Appl Spectrosc 47:1329–1336
89. Noda I (1990) Two-Dimensional Infrared (2D IR) spectroscopy: theory and applications. Appl Spectrosc 44:550–561
90. Noda I (2004) Advances in two-dimensional correlation spectroscopy. Vib Spectrosc 36:143–165
91. Noda I, Ozaki Y (2004) Two-Dimensional correlation spectroscopy—applications in vibrational and optical spectroscopy. Wiley, Chichester
92. Ashton L, Barron LD, Czarnik-Matusewicz B, Hecht L, Hyde J, Blanch EW (2006) Two-dimensional correlation analysis of Raman optical activity data on the á-helix-to-â-sheet transition in poly(L-lysine). Mol Phys 104:1429–1445
93. Czarnik-Matusewicz B, Pilorz S, Ashton L, Blanch EW (2006) Potential pitfalls concerning visualization of the 2D results. J Mol Struct 799:253–258
94. Pazderka T, Kopecky V (2012) Two-dimensional correlation analysis of Raman optical activity—Basic rules and data treatment. Vib Spectrosc 60:193–199
95. Ashton L, Barron LD, Hecht L, Hyde J, Blanch EW (2007) Two-dimensional Raman and Raman optical activity correlation analysis of the á-helix-to-disordered transition in poly(L-glutamic acid). Analyst 132:468–479

96. Barron LD, Blanch EW, Hecht L (2002) Unfolded proteins studied by Raman optical activity. Adv Prot Chem 62:51–90
97. Barron LD, Blanch EW, McColl IH, Syme CD, Hecht L, Nielsen K (2003) Structure and behaviour of proteins, nucleic acids and viruses from vibrational Raman optical activity. Spectroscopy 17:101–126
98. Zhu F, Isaacs NW, Hecht L, Barron LD (2005) Raman optical activity: a tool for protein structure analysis. Structure 13:1409–1419
99. McColl IH, Blanch EW, Gill AC, Rhie AGO, Ritchie MA, Hecht L, Nielsen K, Barron LD (2003) A new perspective on β-sheet structures using vibrational raman optical activity: from poly(L-lysine) to the prion protein. J Am Chem Soc 125:10019–10026
100. Blanch EW, Kasarda DD, Hecht L, Nielsen K, Barron LD (2003) New insight into the solution structures of wheat gluten proteins from Raman optical activity. Biochemistry 42:5665–5673
101. Zhu F, Tranter GE, Isaacs NW, Hecht L, Barron LD (2006) Delineation of protein structure classes from multivariate analysis of protein Raman optical activity data. J Mol Biol 363:19–26
102. Zhu F, Kapitan J, Tranter GE, Pudney PDA, Isaacs NW, Hecht L, Barron LD (2008) Residual structure in disordered peptides and unfolded proteins from multivariate analysis and ab initio simulation of Raman optical activity data. Proteins: Struct Funct Bioinf 70(3):823–833
103. Chruszcz-Lipska K, Blanch EW (2012) In situ analysis of chiral components of pichtae essential oil by means of ROA spectroscopy: experimental and theoretical Raman and ROA spectra of bornyl acetate. J Raman Spectrosc 43:286–293
104. Kinalwa MN, Blanch EW, Doig AJ (2010) Accurate determination of protein secondary structure content from raman and raman optical activity spectra. Anal Chem 82:6347–6349
105. Novak V, Sebestik J, Bour P (2012) Theoretical modeling of the surface-enhanced Raman optical activity. J Chem Theory Comput 8:1714–1720
106. Jensen L, Autschbach J, Krykunov M, Schatz GC (2007) Resonance vibrational Raman optical activity: a time-dependent density functional theory approach. J Chem Phys 127(13):134101
107. Luber S, Neugebauer J, Reiher M (2010) Enhancement and de-enhancement effects in vibrational resonance Raman optical activity. J Chem Phys 132:044113
108. Danecek P, Kapitan J, Baumruk V, Bednarova L, Kopecky JV, Bour P (2007) Anharmonic effects in IR, Raman, and Raman optical activity spectra of alanine and proline zwitterions. J Chem Phys 126:224513

Chapter 5
α-Amino Acids In Water: A Review Of VCD And Roa Spectra

Jan Cz. Dobrowolski, Piotr F.J. Lipiński, Joanna E. Rode and Joanna Sadlej

Abstract An up-to-date perspective on the capability of Vibrational Circular Dichroism (VCD) and Raman Optical Activity (ROA) spectroscopy of the 20 encoded α-amino acids to solve the structural problems appearing in water is provided in this review. VCD and ROA spectroscopy of the α-amino acid molecules reviewed in this chapter refer only to free and single amino acids in water at various pH levels. To well understand the implications of such studies for biology and medicine, we devote one subchapter to the role of free amino acids in an organism's metabolism and physiology. A variety of side chains that modify the α-amino acids character from acidic to basic and from hydrophilic to hydrophobic makes the possible dissociation equilibria in water very complex and dependent on several factors. Therefore, information on dissociation constants, isoelectric points, and hydrophobicity parameters of α-amino acids is summarised in the next subchapter. The conformational variety, complexity of interactions, and entanglement of equilibria makes considering the influence of these factors on VCD and ROA spectra with quantum chemical analysis absolutely necessary. The elements of the theory of chiroptical vibrational spectra and methods enabling solvent simulations are provided in the subsequent subchapter. We report on experimental methods, techniques and tricks for measuring and simulating the VCD and ROA spectra in water at different pH levels. The review is concluded by listing the not-yet-measured amino acids, a presentation of the main challenges for computational methods, and suggestions for the most promising experimental techniques that may be used in future studies.

Keywords Amino acids • Vibrational circular dichroism • Raman optical activity • Water • Experimental • Theoretical • Ab initio • DFT • Quantum mechanical calculation

J. C. Dobrowolski (✉) · J. E. Rode
Industrial Chemistry Research Institute, 8 Rydygiera Street, 01-793, Warsaw, Poland
e-mail: j.dobrowolski@nil.gov.pl

P. J. F. Lipiński
Mossakowski Medical Research Centre
Polish Academy of Sciences
5 Pawinskiego Street, 02-106 Warsaw, Poland

J. C. Dobrowolski · J. Sadlej
National Medicines Institute, 30/34 Chełmska Street, 00-725, Warsaw, Poland

J. Sadlej
Faculty of Chemistry, University of Warsaw, 1 Pasteura Street, 02-093, Warsaw, Poland

M. Baranska (ed.), *Optical Spectroscopy and Computational Methods in Biology and Medicine,* Challenges and Advances in Computational Chemistry and Physics 14,
DOI 10.1007/978-94-007-7832-0_5, © Springer Science+Business Media Dordrecht 2014

5.1 Introduction

Feynman, in his *Lectures on Physics*, invoked a poet's metaphor that "*The whole universe is in a glass of wine*" [1]; it would be organically self-evident, yet maybe more prosaic, if he had added "*and a piece of camembert*". A piece of camembert, looked at closely enough, mirrors the universe's age and the evolution of stars as well as a glass of wine does, but the complexity of phase equilibria, fermentation, the evolution of living creatures, and chirality—the mysterious source of life, would be seen immediately. Also, the presence of mathematics in a piece of camembert is evident not only in Euclidean chirality, displayed by amino acids as a quaternary protein structure, but also in the mathematical logic and combinatorics of primary protein structure, the universal algebra of biochemical processes, and both the discrete and continuous topology of molecular manifolds. Here, we shall focus on the basis of this diversity: the individual natural α-amino acids in a native, aqueous environment.

Usually, the primary importance of amino acids is recognised in their role as the building blocks of proteins fundamental to the cell structure and function of every living organism. However, each free and single α-amino acid at the physiological pH of an organism also plays a fundamental role in maintaining the vital functions of the organism. The Vibrational Circular Dichroism (VCD) and Raman Optical Activity (ROA) spectroscopy of the α-amino acid molecules reviewed in this chapter refer only to free and single amino acids in water at various pH levels. The purpose of this review is to provide an up-to-date perspective on the capability of VCD and ROA to solve the problems of structures for chiral molecules in the solution state. However, to understand the implications of such studies for biology and medicine, we devote subchapter 2 to the role of free amino acids in an organism's metabolism and physiology.

The presence of basic amino and acidic carboxy groups attached to the same carbon atom in an α-amino acid molecule, after dissolution in water, results in the dissociation of these groups and leads to the formation of a zwitterionic form which usually predominates. However, the 20 amino acids encoded by the standard genetic code have a variety of side chains equipped with additional functional groups that modify the amino acid character from acidic to basic and from hydrophilic to hydrophobic. This makes the possible dissociation equilibria of amino acids dissolved in water very complex and dependent on several factors. This is why we summarised information on dissociation constants, isoelectric points, and hydrophobicity parameters of α-amino acids in subchapter 3.

The conformational variety of structures is significant even for small amino acids. Indeed, for neutral valine, 19 structures have been found to be stable, yet, for lysine and arginine, as many as 391 and 520 conformers have been found, respectively [2]. Therefore, even for small amino acids, the interpretation of vibrational spectra requires quantum chemical analysis. The complexity of interactions and equilibria occurring for amino acids in aqueous solutions makes considering the influence of the solution on structure and vibrational spectra absolutely necessary. This is why we provide the reader with subchapter 4, describing elements of the theory of chiroptical vibrational spectra and methods enabling solvent simulations. Having covered the general fundamentals of amino acids and vibrational spectroscopy, the focus of subchapter 5.5 shifts to VCD and ROA results for amino acids.

The Vibrational Circular Dichroism and Raman Optical Activity spectra of amino acids in water have been intensively studied for alanine, but less so for proline. For the other amino acids, there are usually a few papers devoted to their experimental and/or theoretical vibrational chiroptical spectra. Articles comparing the spectra of more than two amino acids are sparse, as this is usually done at the very beginning of the application of chiroptical spectra to the analysis of biomolecules. For eight amino acids, their VCD and ROA spectra were practically unknown. However, at the beginning of 2012, the number of not-yet-studied amino acids was greater by four compounds. For the studied amino acids, we report on experimental methods and techniques as well as tricks for measuring and simulating the VCD and ROA spectra in water at different pH levels.

The review is concluded by listing the not-yet-measured amino acids, a presentation of the main challenges for computational methods, and suggestions for the most promising experimental techniques that may be used in future studies on amino acids in aqueous media.

5.2 Role of Free Amino Acids in an Organism

In humans, the basic set of 20 α-amino acids is categorised into two subgroups: essential and non-essential. The nine essential amino acids cannot be formed in the human body, which is why they have to be ingested with food to maintain the proper function of the organism. On the other hand, there are 11 non-essential amino acids which are synthesised in a simple way from intermediates of the citric acid cycle and other major pathways. The entire set of proteinogenic amino acids can be synthesised by most microorganisms and plants. The biosynthetic routes of most amino acids are greatly interrelated with each other in a high-dimensional network, and also linked to their metabolism and other biochemical processes. There are four metabolic sources of free amino acids:

1. breakdown of ingested proteins
2. breakdown of cellular proteins
3. breakdown of extra-cellular proteins
4. amino acid biosynthesis

5.2.1 Metabolism

The first step of amino acid metabolism is deamination [3]. The -NH_2 group can be transferred onto an acceptor α-ketoacid, which causes the formation of alanine, aspartate, or glutamate. Aspartate enters the urea cycle, while glutamate is easily liberated by oxidative deamination of its amino group in the form of ammonia that also is subject to the same cycle. On the other hand, alanine is an ammonium carrier from the muscles to the liver (glucose-alanine cycle) [4, 5].

The catabolism of amino acid hydrocarbon skeletons can result in the formation of oxaloacetate, a principal metabolite in gluconeogenesis [6, 7]. This is the

case for the so-called glucogenic amino acids (glycine, serine, valine, histidine, arginine, cysteine, proline, alanine, glutamate, glutamine, aspartate, asparagine, and methionine). Lysine and leucine are converted to ketone bodies, and are thus called ketogenic. There is also a group of gluco/ketogenic amino acids (isoleucine, phenylalanine, tryptophan, tyrosine, and threonine) whose hydrocarbon skeleton splits into two fragments that enter the two pathways independently.

5.2.2 Source of Bioactive Molecules

The amino acids are precursors of many bioactive compounds: active one-carbon fragments, porphyrins, purine and pyrimidine bases, hormones, neuromediators, aminoalcohols, biogenic amines, polyamines, carnitine, coenzymes, etc.

Amino acid metabolites often play an important role in signalling. Tyrosine hydroxylated at carbon C3 (DOPA) is converted in the adrenal gland to catecholamine hormones and the neurotransmitters epinephrine (adrenaline) and norepinephrine [8]. Signalling occurs via dopamine, a product of DOPA decarboxylation and a neurotransmitter of great importance, since many psychiatric disorders are associated with the dopamine system.

The decarboxylation of amino acids produces additional biogenic amines beyond dopamine. Among them are γ-aminobutyric acid (GABA) (from glutamate), a principal inhibitory neurotransmitter; histamine (from histidine), a neurotransmitter and immune response molecule; and serotonin (from tryptophan via 5-hydroxytryptophan), another neurotransmitter involved in psychiatric disorders [9–12].

Other hormones related to tyrosine are triiodothyronine and thyroxine that are biosynthesised via the modification of tyrosine residues in thyroglobulin.

Not only hormones are produced from amino acids. A local signalling molecule, nitric oxide, is involved in, inter alia, vessel homeostasis and adaptive vasodilatation, and is made from arginine [13].

Amino acids are also involved in the energetics of cells and organisms. Carnitine, which transports fatty acids from the cytosol into the mitochondria during the breakdown of lipids, is synthesised from lysine and methionine, while creatine, part of the phosphocreatine system used to help cells (muscles) meet increased energy demands during anaerobic effort, is made of L-arginine, glycine, and L-methionine [14, 15].

L-methionine takes part in transmethylation reactions via its metabolite S-adenozylomethionine (SAM) with a highly reactive methyl group [16]. SAM plays a role in the biosynthesis of polyamines, as well as of dopamine and serotonin. Another source of active one-carbon fragments is serine, which donates part of its skeleton to tetrahydrofolate (THF), forming 5-methylotetrahydrofolate (5-MTHF), which subsequently passes its methyl group on to homocysteine to yield methionine [17]. The entire process is called the active methyl cycle. 5-MTHF participates also in the synthesis of purines and thymidine [18].

Amino acids are necessary for the formation of some important structural molecules. The sulphur of cysteine is oxidised, liberated, and conjugated to ATP to form an active sulphate, a coenzyme in the synthesis of glycosaminoglycan, which builds

connective tissue. Serine builds cell membranes as a phospholipid, i.e. phosphatidylserine. Porphyrins (e.g. heme, a cofactor of haemoglobin) are formed, inter alia, from glycine [19].

Although the view of amino acid function is generally dominated by their role in humans (or higher animals), it should be stressed that they are also important molecules in bacteria, plants, and animals. Their role as signalling molecules (or precursors thereof) seems to be most interesting, especially in an evolutionary context.

An emerging and interesting field of research is focused on the D-amino acids, as the role of these molecules is now being recognised. In bacteria and some invertebrates, an essential role is played by D-alanine [20–22]. In these organisms, alanine racemases catalyse the interconversion of D- and L-alanine. In bacteria, D-alanine is incorporated into bacterial cell walls [20, 21], while in sea invertebrates, it is one of the major compounds responsible for intracellular osmotic regulation [22]. On the other hand, fish species eating crustaceans and molluscs containing D-amino acids use oxidases to catalyse the decomposition of these compounds. Free D-alanine has also been found in the mammalian brain, liver, kidney, blood, and urine [23]. An increase in D-alanine in the brain had been suggested to be associated with the progression of Alzheimer's disease [24, 25] and in plasma with renal disease [26]. It has been demonstrated that, in addition to D-alanine, D-aspartic acid, D-serine, D-leucine, and D-proline are also present in the mouse brain and their degradation by D-amino acid oxidase has been investigated [27].

Another recognised D-amino acid is D-aspartic acid, present in invertebrate and vertebrate neuroendocrine tissues, where it carries out important physiological functions and is implicated in nervous system development [28]. It was recently shown that D-aspartic acid is a novel endogenous neurotransmitter in distantly related animals, and is present in high concentrations in the synaptic vesicles of axon terminals [28, 29]. The synthesis of D-aspartic acid occurs in neurons by conversion of L- to D-aspartic acid via D-aspartate racemase, which is the main source of neuronal D-aspartic acid [30]. Another example of an important D-amino acid is D-serine, which modulates neurotransmission [31].

For some other important facts about the role of free amino acids in organisms the reader is referred to the literature and biochemical textbooks [18, 32, 33].

5.3 Amino Acids in Aqueous Solutions

5.3.1 Acid Base Equilibria

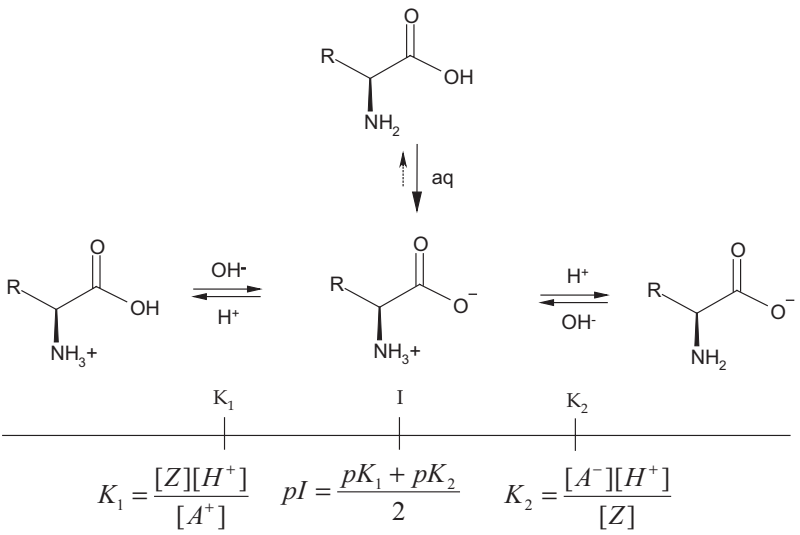
A model α-amino acid molecule is composed of an aliphatic skeleton with one carboxy and one amino group which may dissociate in water. Such a simple case means that the neutral, zwitterion, and anion or cation of an amino acid may coexist, depending on the acidity of the surrounding media and the acid-base characteristics of a given amino acid (Scheme 1). The equilibria become more and more complicated with every additional functional group present in the molecule. An additional terminal acidic

COOH group is found in aspartic and glutamic acids, an additional terminal amide group occurs in asparagine and glutamine, an extra terminal amine group is present in lysine, additional aliphatic OH groups appear in serine and threonine, while phenolic OH is present in tyrosine. Moreover, an additional NH group may be incorporated into a bicyclic aromatic system, such as in tryptophan, or in an apparently simple five-membered imidazole ring (allowing for additional tautomeric equilibria), as in histidine. Obviously, the SH, SCH_3, and aromatic ring groups, as in cysteine, methionine, and phenylalanine, respectively, change the complexity of equilibria in water as well. Furthermore, some parts of amino acid molecules are hydrophobic and force specific conformations to be populated in water. Last but not least, amino acid molecules are flexible molecules and usually exhibit a fascinatingly rich conformational landscape. All these complex problems must be taken into account when one aims to measure the (chiroptical) vibrational spectra of single, free amino acid molecules in aqueous solution and/or to model them by quantum chemical methods.

Let us now provide some basic facts about α-amino acid molecules in water, their acid-base equilibria, hydrophobic-hydrophilic character and solubility. First, note that acetic acid and ammonia, model molecules for the two functional groups of amino acids, exhibit dissociation constants of similar values:

$$CH_3COOH + H_2O \leftrightarrow CH_3COO^- + H_3O^+ \qquad K_a \approx 1.8\cdot10^{-5}$$

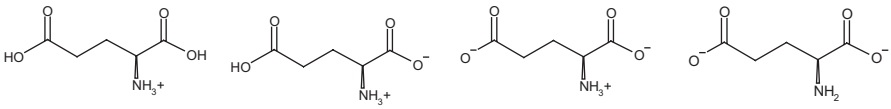
$$NH_3 + H_2O \leftrightarrow NH_4^+ + OH^- \qquad K_b \approx 1.8\cdot10^{-5}$$



Scheme 1 Scheme of the simplest acid base equilibria occuring for an α amino acid molecule in aqueous solution. I—isoelectric point, R side chain

This makes α-amino acid molecules in water capable of forming a dipolar ion by the transfer of a proton from the carboxy to the amino group. Then, the COO^- and NH_3^+ ionic moieties are simultaneously present in the molecule, forming a zwit-

terion (from the German *Zwitter*, meaning hybrid). The maximum concentration of the stable zwitterion is characteristic of each individual amino acid and occurs at the isoelectric point *I*, which is determined by the formula $pI = ½(pK_1 + pK_2)$. Although, at the isoelectric point, the population of the zwitterion is maximum, the presence of the other amino acid forms is not always negligible. However, in water, the non-ionised neutral form of amino acids is a very minor species.



Scheme 2 Four forms of glutamic acid considered in protolytic equilibria in water.

A change in pH from the isoelectric point to more basic or more acidic conditions increases the population of anions or cations, respectively. Moreover, if an additional dissociating functional group is present in the amino acid molecule, it gives rise to the formation of more complex equilibria. On the other hand, if an additional dissociating group is absent, the *pI* value of the proteinogenic α-amino acid at 25 °C varies between 5.4 and 6.3 (Table 5.1). Below or above the pH determined by pK_1 and pK_2 (Table 5.1), the cation or anion are the prevailing forms, respectively. Although the dissociation constants of the amino groups in the α position are only slightly different from that of ammonia ($pK_b = 9.24$), this is not the case with the juxtaposition of the carboxylic group with that of acetic acid. The presence of the α-amine group causes the pK_a to sharply increase to ca. 4.75. In amino acids bearing an R substituent with ionisable groups, the double cations (histidine, lysine, and arginine) or double anions (aspartic acid, glutamic acid, but also cysteine and tyrosine, Table 5.1) can prevail at their specific pH levels. For example, in a recent study on the protolytic equilibria of L-glutamine and L-glutamic acid, three forms of glutamine and four forms of glutamic acid were found to be present in the pH range from 0 to 12 (Scheme 2) [36]. Similarly, four forms of aspartic acid were considered to occur at equilibrium in acetonitrile [37].

5.3.2 *Solubility of Amino Acids in Water*

Measuring amino acids in water requires the formulation of adequate solutions; however, this is very difficult for many of them. Although α-amino acids in their native conditions are dissolved in various physiological media, their solubilities in water (Table 5.1) range from very large (proline, glycine, and alanine) through moderate (valine, threonine, isoleucine, asparagine, and glutamine) to low or very low (aspartic and glutamic acids, lysine, and tyrosine). A change in the surrounding acidity or basicity increases the solubility of the ionic forms of amino acids. Also, the presence of other water soluble molecules brings about an increase in solubility. Nevertheless, for quite a number of amino acids, vibrational spectroscopy measurements in water are very challenging and require tricks [38]. Because of their high polarity, amino acids are insoluble in inert and slightly polar solvents and barely soluble in ethanol [34].

Table 5.1 α-Amino acids: dissociation constants and isoelectric points at 25 °C and solubility in water at different temperatures (g/100 g). [34]

Amino	Dissociation constants					Solubility in water, temp. °C				
Acid	pK_1	pK_2	pK_3	pK_4	pI	0	25	50	75	100
Alanine	2.34	9.69			6.0	12.73	16.51	21.79	28.51	37.30
Arginine	2.18	9.09	12.60		10.8		18.1 [35]			
Asparagine	2.02	8.80			5.4		2.51 [35]			
Aspartic acid	1.88	3.65	9.60		2.8	0.209	0.500	1.199	2.875	6.893
Cysteine	1.71	8.35	10.66		5.0		*vg* [35]			
Cystine	1.04	2.10	8.02	8.71	5.1		0.011 [35]			
Glutamine	2.17	9.13			5.7		4.2 [35]			
Glutamic acid	2.19	4.25	9.67		3.2	0.341	0.843	2.186	5.532	14.00
Glycine	2.34	9.60			6.0	14.18	24.99	39.10	54.39	67.17
Histidine	1.80	5.99	9.07		7.5	–	4.29	–	–	–
4-Hydroxy-proline	1.82	9.65			5.7	28.86	36.11	45.18	51.67	–
Isoleucine	2.36	9.68			6.0	3.791	4.117	4.818	6.076	8.255
Leucine	2.36	9.60			6.0	2.270	2.19	2.66	3.823	5.638
Lysine	2.20	8.90	10.28		9.6		0.58 [35]			
Methionine	2.28	9.21			5.7	1.818	3.381	6.070	10.52	17.60
Phenylala-nine	1.83	9.13			5.5	1.983	2.965	4.431	6.624	9.900
Proline	1.99	10.60			6.3	127.4	162.3	206.7	239.0	–
Serine	2.21	9.15			5.7	2.204	5.023	10.34	19.21	32.24
Threonine	2.15	9.12			5.6		9.81 [35]			
Tryptophan	2.38	9.39			5.9	0.823	1.136	1.706	2.795	4.987
Tyrosine	2.20	9.11	10.07		5.7	0.020	0.045	0.105	0.244	0.565
Valine	2.32	9.62			6.0	8.34	8.85	9.62	10.24	–

vg very good

5.3.3 Hydrophilicity vs. Hydrophobicity

The other factor accounting for both the solubility of amino acids and the complexity of the acid-base equilibria of amino acids in water solution is the hydrophobicity vs. hydrophilicity of the side chain. This property of amino acids is extremely important in biology. Hydrophobic forces drive the formation of secondary, tertiary, and quaternary protein structures. The character of side chains plays an important role in protein location at interfaces or in the transmembrane space. It also influences the behaviour of individual amino acids in solution as well as the structure of water [39].

According to the IUPAC Gold Book definition [40], hydrophobicity is the association of non-polar groups or molecules in an aqueous environment which arises from the tendency of water to exclude non-polar molecules. The origin of the effect lies in the disruption of the hydrogen bond network between molecules of water

after the introduction of a non-polar compound which cannot form hydrogen bonds and thus replaces the disrupted ones.

Apart from the qualitative description of amino acids as hydrophobic or hydrophilic, there have been numerous attempts at quantitatively characterising the hydrophobicity of amino acids by means of both experimental and theoretical methods. An exhaustive review on different scales is provided by Biswas et al. [41]. The derived hydrophobicity scales were divided by the authors according to the type of studied solutes (simple amino acids, side chain analogues, their derivatives, or peptides) and the method used (partitioning experiments, chromatography, site-directed mutagenesis, accessible surface area calculations, and the measurement of physical properties).

Depending on the method used, hydrophobicity is expressed as a different experimental observable. In partitioning approaches, one uses the free energy of transfer between phases, while in chromatographic studies, hydrophobicity is related to retention factors. There have also been attempts to quantify this property as a function of some physical property, e.g. surface tension, transition temperature, solvation energy, and partial molar heat capacity [41].

Table 5.2 presents amino acids ranked according to different hydrophobicity scales. Levitt's scale is based on the solubility of single amino acids in water and ethanol and the derived free energy of transfer between the two phases [42]. Parker et al. used for his hydrophobicity scale the HPLC parameters of some synthetic oligopeptides where two residues were sequentially substituted with different amino acids [43]. Radzicka and Wolfenden determined the distribution coefficients of solutes related to amino acid side chains (e.g. toluene for phenylalanine, methanol for serine, isobutene for leucine, etc.) in water and wet cyclohexane [44]. Kyte and Doolittle combined various experimental observations in the literature [45]; their hydropathy scale was based on water vapour transfer free energies and the interior-exterior distribution of amino acid side chains determined by Chothia [46]. Janin [47] and Rose et al. [48] examine the tendency of amino acid side chains to be found inside (hydrophobicity) or outside (hydrophilicity) some proteins with known 3-D structures. One of the newest and most interesting scales is the one proposed by Moret and Zebende [49]. They used variations in the accessible surface area determined for amino acids in short protein fragments randomly extracted from PDB proteins as a measure of hydrophobicity.

The scales based on different assumptions differ, as shown in Table 5.2. Particular compounds are not ranked in the same order according to different scales. In many cases, there is no correlation between scales, whether theoretical or experimental, although there are some regularities. For example, charged amino acids are generally more polar [49]. It is very important to stress that most measurements have been performed on peptides, and only rarely on single amino acids (or their derivatives). This is why the length, structure, and kind of peptide used, the neighbourhood of the studied amino acid, etc. substantially influence the obtained values. Another non-trivial issue is the selection of proper experimental

Table 5.2 α-Amino acids ordered from hydrophilic to hydrophobic according to different scales

Levitt [42]	Parker et al. [43]	Radzicka & Wolfenden [44][a]	Kyte & Doolittle [45]	Chothia [46]	Janin [47]	Rose et al. [48]	Moret & Zebende [49]
Arg	Asp	Arg	Arg	His	Lys	Lys	Lys
Lys	Glu	Asp	Lys	Arg	Arg	Glu	Arg
Glu	Asn	Glu	Asp	Lys	Glu	Asp	Asp
Asp	Ser	Asn	Glu	Gln	Gln	Gln	Glu
Ser	Gln	Lys	Asn	Asn	Asp	Asn	Ser
Gln	Lys	Gln	Gln	Asp	Asn	Arg	Gln
Asn	Gly	His	His	Tyr	Tyr	Pro	Asn
Gly	Thr	Ser	Pro	Glu	Pro	Ser	Pro
Thr	Arg	Thr	Tyr	Pro	Thr	Thr	Thr
His	Pro	Pro	Trp	Ser	His	Gly	His
Ala	His	Tyr	Ser	Thr	Ser	Ala	Gly
Cys	Ala	Gly	Thr	Trp	Trp	Tyr	Ala
Met	Cys	Cys	Gly	Gly	Gly	His	Trp
Pro	Tyr	Ala	Ala	Ala	Ala	Trp	Leu
Val	Val	Trp	Met	Met	Met	Met	Phe
Ile	Met	Met	Cys	Leu	Phe	Leu	Met
Leu	Ile	Phe	Phe	Phe	Leu	Val	Tyr
Tyr	Phe	Val	Leu	Cys	Val	Ile	Ile
Phe	Leu	Leu	Val	Val	Ile	Phe	Val
Trp	Trp		Ile	Ile	Cys	Cys	Cys

[a] Isoleucine not included in the analysis

settings. For example, in partitioning methods, it is the choice of the phase to be used, i.e. liquid, vapour, or micellar, as well as the components of the immiscible solvent system (e.g. 1-octanol/water; cyclohexanol/water etc.) that are the most important.

Even though hydrophobicity parameters are assigned to individual amino acids, they are usually derived from peptide properties, and the hydrophobicity scales are first and foremost used in a range of fields related to proteins. The standard and most common application of such scales is in identifying surface-exposed regions as well as transmembrane regions and for the prediction of protein secondary structure [50]. Another important application of these scales is in studying the structural homology between proteins and as an aid in sequence alignment [51]. Finally, some were developed for peptide QSAR (Quantitative Structure–Activity Relationship) and QSPR (Quantitative Structure-Property-Relationships) purposes [52].

It is certain that hydrophobicity has an important role for single amino acids in solution; however, the derived scales, although quite useful for the above mentioned applications, give only a rough insight into the meaning of spectroscopy measurements of individual amino acids in aqueous solution.

5.4 Basis of Molecular Activity in Vibrational Spectra

5.4.1 General Considerations

There are many methods which can be used to determine the structure and to understand the function of biomolecules. The methods of choice for biomolecules that can be crystallised are X-ray and neutron diffraction. For several smaller and more flexible biomolecules, such as some amino acids and small peptides, this is hardly possible. In these cases, NMR techniques, using the nuclear Overhauser effect, have been used to supplement the diffraction methods and to solve the structure of biomolecules in the folding state in solution.

Optical spectroscopy methods have also provided an effective way of investigating molecular structure. These techniques are characterised by lower resolution than X-ray crystallography (as the wavelength of infrared or visible light is much longer than typical molecular sizes) for the registration of signals from individual conformers. Increased conformational sensitivity can be achieved with polarised techniques by assessing absorption or scattering of circularly polarised light by chiral systems.

In recent years, a renaissance in the use of chiroptical spectroscopic methods has taken place. Books and reviews are available on this topic [53–68]. Vibrational Optical Activity (VOA) encompasses two main forms of chiroptical spectroscopy: (1) vibrational circular dichroism (VCD) and (2) vibrational Raman optical activity (ROA). These methods are powerful techniques in stereochemical analysis (conformation as well as configuration), and in determining the absolute configuration (AC) of molecules. Normal and chiral spectroscopies are highly structure-sensitive. Many changes have been observed in the spectra as a function of solvent polarity, pH, and temperature. Typical VOA (VCD/ROA) intensities are four or five orders of magnitude smaller than those of the parent IR/Raman methods. Nevertheless, use of VCD/ROA spectra combined with quantum-chemical calculations is a method of choice for the assignment of AC instead of intense but much less selective optical rotatory dispersion (ORD) or electronic circular dichroism (ECD) spectra.

5.4.2 Chiroptical Methods

The chiroptical spectroscopic methods that are in wide current use include: 1. electronic circular dichroism (ECD), 2. optical rotatory dispersion (ORD), 3. vibrational circular dichroism (VCD), and 4. vibrational Raman optical activity (ROA). Chiroptical VCD or ROA spectroscopies probe molecular vibrations to measure the differential absorption (IR) or scattering (from NIR to UV) of right- and left-circularly polarised light. In contrast to the ECD technique, these two methods can be applied to molecules lacking chromophores. Moreover, the numerous vibrational modes give more data for comparison with the simulated spectra than a few broad electronic spectroscopy bands.

5.4.2.1 Vibrational Circular Dichroism (VCD)

VCD is an extension of the electronic CD technique into the IR region of the electromagnetic field. The spectrum usually consists of numerous bands in the 800–2000 cm^{-1} region, which is the advantage of this method. Nevertheless, a disadvantage of VCD in comparison to ECD is the intensity, about four to five orders of magnitude lower than that of the parent IR vibrational absorption spectra vs. ca. three orders in the parent UV-Vis spectra [69]. This explains the difficulties with observation of VCD spectra before the era of Fourier-transform IR spectrometers. This method has become popular for elucidating AC by correlation between the experimental and calculated VCD spectra. Usually, only fundamental vibrational transitions are considered in the VCD spectra. As for the IR absorption spectra, the VCD band frequency yields information on the energy of a particular vibrational mode (the frequencies), but the VCD band intensity can take either a positive or negative sign.

The first VCD measurements were published in 1972 by Dudley et al. for cholesteric mesophase [70] and by Chabay, Holzwarth and Hsu for tartrate salt and inorganic salts [71–73]. The theoretical fundament for VCD theory was established by Barron, Buckingham, and Raab in 1975 [74, 75]. Nafie et al. presented good quality VCD spectra of a number of chiral molecules in the liquid phase in 1973 [76]. The first FTIR-VCD spectrometer at Syracuse University was designed and built in 1979 [77], while the first commercially available VCD spectrometers appeared on the market in the mid-1990s [55].

Experimental VCD measurements are usually carried out in IR transmitting solvents or, in high concentrations, in light or heavy water. The pioneering work by Schlosser et al. [78] and Henderson and Polavarapu [79] opened new possibilities for experimental investigations using low-temperature matrix isolation vibrational circular dichroism spectra (MI-VCD). Recently, a contemporary, experimental matrix isolation-VCD paper by Tarczay et al. for (R)-2-amino-1-propanol [80] has been published, and the VCD spectra of Ac-Gly-NHMe and Ac-L-Ala-NHMe were registered in matrices [81]. The VCD signals of water molecules interacting with peptides were detected [82, 83].

Now, we briefly describe the basic terms involved in the calculations of VCD spectra. For a more detailed description, we refer the interested reader to some excellent reviews [55, 61, 68, 84, 85]. At the molecular level, for each mode, the intensity of the VA band is proportional to the absolute square of the electric-dipole transition moment (EDTM) of the molecule called the dipole strength of the transition. For the *i*-th mode, this is denoted by $D_{g1,g0}(i)$, expressing the transition from the ground electronic (*g*) and ground vibrational state (*0*) (described by the Ψ_{g0} wave function) to the ground electronic (*g*) and vibrationally excited state (*1*) (described by the Ψ_{g1} wave function). The dipole strength is always positive as an absolute square. The corresponding intensity of the VCD band of the *i*-th mode is connected to the rotatory strength $R_{g1,g0}(i)$: the imaginary part of the scalar product of EDTM and the magnetic-dipole transition moment (MDTM).

At the molecular level, these two quantities for the *i*-th vibrational mode are defined as follows:

$$D_{g1,g0}(i) = \left|\left\langle \Psi_{g0} \left| \mu_{el} \right| \Psi_{g1} \right\rangle\right|^2 = E^2_{g1,g0}(i) \tag{5.1}$$

$$R_{g1,g0}(i) = Im\left[\left\langle \Psi_{g0} \left| \mu_{el} \right| \Psi_{g1} \right\rangle \left\langle \Psi_{g1} \left| \mu_{magn} \right| \Psi_{g0} \right\rangle\right] = E_{g1,g0}(i) Im\left[M_{g1,g0}(i)\right] \tag{5.2}$$

The sign of the rotatory strength of the *i*-th mode is thus determined by the cosine of the angle ξ between the EDTM and MDTM vectors, $E_{g1,g0}(i)$ and $M_{g1,g0}(i)$.

When $\xi < 90°$ the sign of the VCD intensity is positive ($R(i) > 0$), whereas when $\xi > 90°$ it is negative ($R(i) < 0$). In achiral molecules, the two vectors (generally different from zero) are perpendicular and the scalar product is equal to zero ($R(i) = 0$). For ξ close to 90°, even small perturbations produced by the solvent, interactions, molecular conformation, or, in computations, by an inadequate theory level or basis set, may change ξ across the 90° and thus may induce an erratic change of the sign. This problem will be discussed later on.

Electric dipole derivatives are analysed with so-called atomic polar tensors (APTs), which are defined as molecular dipole derivatives with respect to the Cartesian coordinates of nuclei at the equilibrium geometry. The so-called atomic axial tensor (AATs) has been introduced, for which entries are derivatives of the ground state molecular magnetic moment with respect to the velocity of nuclei. The magnetic-dipole moment operators in the rotatory strength tensor are dependent on the arbitrary gauge origin. Therefore, although the EDTM causes no problems in computations with a finite atomic orbital basis set, the response properties involving MDTM are origin dependent, i.e., they may change if an origin shift is applied to the coordinates of the molecule. The Gauge Invariant (or Including) Atomic Orbital (GIAO) basis, where each individual basis function is made to depend explicitly on the external magnetic field induction, eliminates this origin dependence even for a finite basis set and it is now the method of choice [86].

The most successful realisation of the above algorithm is the perturbative Stephens approach [87, 88] within the Born-Oppenheimer approximation, the double harmonic approach in the equations of APT and AAT, and the use of harmonic approximation for frequency calculations. This approach has been implemented in several computer programs [89–92] and has been reviewed a number of times in recent years [55, 58, 61, 68, 84, 85]. This approach has been recently modified by Bouř et al. [93] and by Coriani et al. [94], who have introduced an alternative time-saving formula for the AAT tensor computations.

In recent years, considerable progress has been reported in the Coupled Clusters (CC) description and calculations of the electronic structure and the molecular properties of molecules at higher levels of the electron correlation effect. For the CC method, the expression for magnetic-dipole vibrational transition moment must be evaluated with consideration of the differing left- and right-hand coupled-cluster functions [95]. There are now two available quantum chemistry program packages,

DALTON [91] and PSI [96, 97] developed by the Ruud and Crawford laboratories, respectively, capable of computing the chiroptical response and circular dichroism spectra using the CC linear-response method. This method provides the best comparison to experimental gas-phase VCD spectra.

Practical simulation of VCD spectra A recent increase in the accessibility of the VCD spectrometers resulted in intense development of quantum chemistry programs calculating VCD intensities directly from first principles. In the following subsection, we will briefly discuss the practical aspects of the computational methods used for the calculation of the VCD spectra.

Chiroptical properties are sensitive to the approximation made in electronic structure calculations. Therefore, it is advisable to use the most accurate method that can be afforded for a given molecule. The study of VCD of a chiral molecule requires several steps and begins with the determination of the molecular structure.

1. The first step is to determine local minimum structures by a systematic conformational search or by molecular dynamics (MD) or by Monte Carlo (MC) simulations. For the given optimised geometry obtained by quantum mechanical calculations, the harmonic force field and normal coordinates are then found.
2. Calculation of the VCD spectra, within the quantum mechanical approaches, as a second step, can be performed at different levels of theory. Currently, most studies of VCD spectra employ the Density Functional Theory (DFT) approach. These computations give reasonably accurate results at a moderate computational cost. Of course, the simulated VCD spectra are sensitive to the DFT functionals and basis sets [98]. Hybrid functionals with a large fraction of Hartree-Fock exchange (50%) are often employed for these calculations. Having the structures and their relative energies for the different conformers, the VCD spectrum has to be calculated for individual conformers. This means that the calculations of the derivatives of the electric-dipole transition moments (as for IR absorption spectra) must be followed by the calculations of the magnetic-dipole moments. The final rotational strengths are computed as products of both transition moments.
3. The third step is to obtain the final spectrum; thus, the rotational strength of the spectrum of each conformer has to be scaled by a factor that reflects the relative Boltzmann population N_i of a given conformer at a given temperature T, calculated with the Gibbs free energy difference ΔG_i of the i-th conformer with respect to the most stable conformer. Boltzmann averaging gives spectra that can be compared with experimental results.

5.4.2.2 Vibrational Raman Optical Activity (ROA)

ROA is described by means of the absolute difference between the intensities of the (inelastic) scattering of right (R) and left (L) circularly polarised incident light by

a chiral sample. It was defined by Atkins and Barron in 1969 [99] as a dimensionless circular intensity difference (CID), Δ, now called incident circular polarisation ROA (ICP ROA),

$$\Delta = (I^R - I^L)/(I^R + I^L) \tag{5.3}$$

where I^R and I^L are the scattered intensities in right- and left-circularly polarised incident light. Barron and Buckingham gave the theoretical background for this effect [100]. The first experimental ROA spectra were published in 1973 [101] and were confirmed by Hug et al. in 1975 [102].

The intensity of ROA spectra is weak, ca. 10^{-4} of the parent vibrational Raman scattering. Experimental spectra can be registered using different scattering angle: forward (0°), or right-angle (90°), or backward (180°). The conventional Raman scattering intensities are identical in the forward and backward modes of operation. For three scattering geometries there are three basic circular polarisation schemes: incident circular polarisation (ICP), scattered circular polarisation (SCP), and dual circular polarisation (DCP), the last one being a combination of ICP and SCP ROA. In the ICP scheme, the incident laser is modulated between R- and L-circular polarisation states and the intensity is measured at a fixed linear or unpolarised radiation, while in the SCP scheme, fixed linear or unpolarised incident laser radiation is used and the difference in the R- and L-circularly polarised scattered light is measured. Within the far-from-resonance approximation, the information obtained with ICP and SCP schemes are the same [103]. The expressions for specific experimental schemes for ROA can be found in the literature [104]. The best experimental choice for most ROA registrations, particularly for studies of biomolecules in aqueous solution, is the backward mode. Absolute scattering cross-sections are available for molecules in the gas-phase; however, when ROA spectra are measured for liquids, the intensity of ROA is modified [53].

Within the Born-Oppenheimer approximation, (where only the linear contribution in the Taylor series of polarisabilities about the equilibrium geometry contributes significantly), double harmonic approximation and the far-from-resonance theory, as well as within the Placzek polarisability theory of the Raman spectra the ROA intensity differences for the SCP backscattered mode can be written in terms of the normal coordinate derivatives of three polarisability tensors: the electric dipole-electric dipole polarisability tensor, the electric dipole-magnetic dipole polarisability tensor, and the electric dipole-electric quadrupole polarisability tensor, as follows [105]:

$$I^R + I^L \propto \frac{\left(\bar{\nu}_{in} - \bar{\nu}_i\right)^4}{1 - \exp\left(-hc\bar{\nu}_i / k_B T\right)}\left[45\alpha_i^2 + 7\beta_i^2\right] \text{ Raman intensities} \tag{5.4}$$

$$I^R - I^L \propto \frac{\left(\bar{\nu}_{in} - \bar{\nu}_i\right)^4}{1 - \exp\left(-hc\bar{\nu}_i / k_B T\right)}\left[12\beta_{G_i}^2 + 4\beta_{A_i}^2\right] \text{ ROA intensity differences} \tag{5.5}$$

where

$$\alpha_i^2 = \frac{1}{9}\left(\frac{\partial \alpha_{\alpha\alpha}}{\partial Q_i}\right)_0 \left(\frac{\partial \alpha_{\beta\beta}}{\partial Q_i}\right)_0$$

$$\beta_i^2 = \frac{1}{2}\left[3\left(\frac{\partial \alpha_{\alpha\beta}}{\partial Q_i}\right)_0 \left(\frac{\partial \alpha_{\alpha\beta}}{\partial Q_i}\right)_0 - \left(\frac{\partial \alpha_{\alpha\alpha}}{\partial Q_i}\right)_0 \left(\frac{\partial \alpha_{\beta\beta}}{\partial Q_i}\right)_0\right]$$

$$\beta_{G_i}^2 = \frac{1}{2}\left[3\left(\frac{\partial \alpha_{\alpha\beta}}{\partial Q_i}\right)_0 \left(\frac{\partial G'_{\alpha\beta}}{\partial Q_i}\right)_0 - \left(\frac{\partial \alpha_{\alpha\alpha}}{\partial Q_i}\right)_0 \left(\frac{\partial G'_{\beta\beta}}{\partial Q_i}\right)_0\right]$$

$$\beta_{A_i}^2 = \frac{\omega}{2}\left(\frac{\partial \alpha_{\alpha\beta}}{\partial Q_i}\right)_0 \left(\frac{\varepsilon_{\alpha\gamma\delta}\partial A_{\gamma_i\delta\beta}}{\partial Q_i}\right)_0$$

In expressions for Raman and ROA observables, eqs. (5.4) and (5.5), symbols α^2_i and β^2_i are the Raman invariants, β^2_{Gi} and β^2_{Ai} are the ROA invariants; the Q_i is the normal mode of the *i*-th vibration, and the three $\alpha^2_{\alpha\beta,}$ $G'_{\alpha\beta}$ and $A_{\alpha\beta\gamma}$ polarisability derivatives are the frequency dependent electric dipole-electric dipole, electric dipole-magnetic dipole, and electric dipole-electric-quadrupole polarisabilities, respectively, used to form the three ROA tensor invariants; $\varepsilon_{\alpha\beta\gamma}$ is the element of the Levi-Civita antisymmetric tensor, ν_{in} is the frequency of the incident light, ν_i is the frequency corresponding to the vibrational excitation of the *i*-th vibrational mode, and k_B is the Boltzmann constant. All derivatives are evaluated at the molecular equilibrium geometry, indicated by the subscript "*0*". The electric dipole-electric dipole polarisability derivatives $\alpha^2_{\alpha\beta}$ are used to form two Raman tensor invariants. $\alpha^2_{\alpha\beta,}$ $G'_{\alpha\beta}$ and $A_{\alpha\beta\gamma}$ are defined in terms of sum-over-states expressions for exact wave functions; see Ref. [104]. The subscript Greek letters in these equations refer to Cartesian componets and the Einstein summation convention is used. In the limit of the static field (ν_{in}) the G' tensor will vanish.

Practical simulation of ROA spectra The *ab initio* computation of ROA requires the differentiation of the optical tensors, **α**, **G'** and **A** with respect to the nuclear coordinates. The first *ab initio* spectra were computed using the Hartree-Fock approach by Polavarapu et al. in 1990 [106]. In these calculations, the geometrical derivatives of the different electronic tensors were obtained by numerical differentiation of the electronic tensors with respect to the Cartesian displacement of the nuclei. Also, the conventional basis set was used and, as a consequence, the calculated tensor invariants were gauge origin dependent. The numerical procedure used for the nuclear displacement derivatives of polarisabilities restricted the size of molecules that can be studied. In 1994, Helgaker et al. presented the GIAO calculations of ROA at the HF and multiconfigurational SCF (MC SCF) levels [107]. An important step in the development of ROA calculations was made by using, for the first time, the analytical DFT approach, with a favourable cost-accuracy ratio, and by the implementation of linear response theory and GIAO by Ruud et al. [108]. The convergence of the ROA parameters with the basis set size was shown to be

slower than that of conventional Raman intensities, and at least the d-aug-cc-pVTZ basis set should be used to avoid an error larger than 10% with respect to the basis set limit [55]. The important step in the development of ROA calculations was made using the DFT approach and by the implementation of linear response theory [108]. As for VCD, the more recent advances in ROA calculations were made by the introduction of higher-level electron-correlation methods such as the coupled-cluster approach (CC) [95]. The fully consistent analytic implementations of ROA have started to appear only very recently; see Refs. [104, 105] for a review of theoretical approaches for computing ROA spectra.

The ROA study of a chiral molecule requires several steps, beginning from the determination of the molecular structure. The intensity differences are dependent on the quality of the force field. For fully analytical calculations, Cheeseman and Frisch [105] proposed two alternatives: a "one-step" and a "two-steps" procedure. In the former, the force field and Raman/ROA tensor invariants are computed at the same time using the same level of theory. In the "two-steps" procedure, the force field and Raman/ROA tensor invariants are computed in separate steps, allowing for the use of different levels of theory at each step.

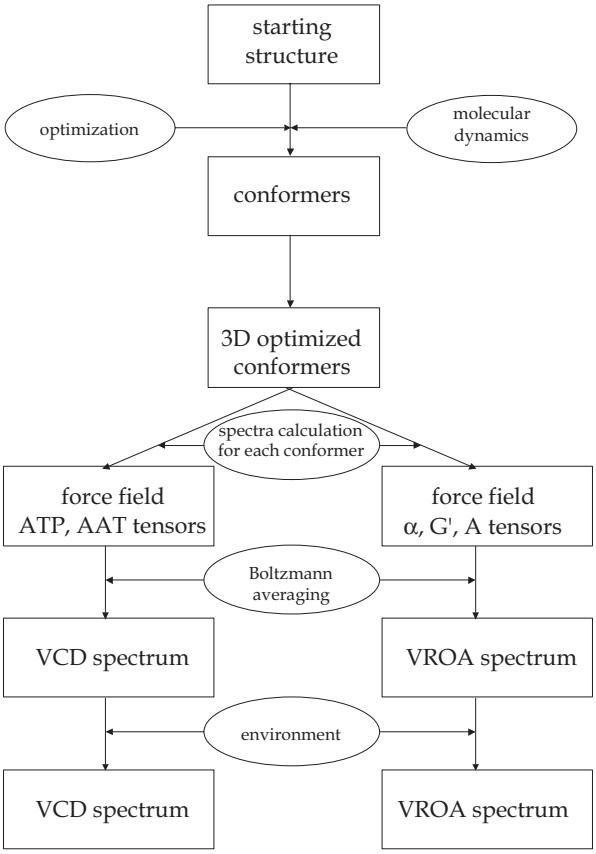
To summarise:

1. As for VCD, the first step is to determine local minimum structures from a systematic conformational search, molecular dynamics (MD), or Monte Carlo (MC) simulations. The harmonic force field and normal coordinates are found for a given optimised geometry.
2. Calculation of the ROA spectra in the frame of the quantum chemistry can be performed at various levels of theory. Currently, most studies of ROA spectra use the DFT approach. Raman and ROA intensities are known to be very sensitive to the basis set, as tensor invariants require the diffuse augmentation of the basis set [105].
3. Finally, to obtain the ROA spectrum that can be compared with the experimental one, the intensity of each conformer's spectrum has to be scaled by a factor reflecting the relative Boltzmann population N_i of a given conformer in the equilibrium conformer mixture at a given temperature T.

These calculations are recommended to both vibrational chiroptical methods to carefully investigate the problems [59]. The current status of the field is highlighted by Barron and Buckingham [109]. A schematic representation of the computations of chiroptical VCD and ROA spectra is presented in Scheme 3. Various quantum chemical codes which have implemented chiroptical properties are available now: Gaussian 09 [89], ADF [90], DALTON [91], CADPACK [92], PSI [96, 97], and MoViPac [110].

Recently, experimental and computational optical vibrational activity have become one of the mainstream research domains in physical chemistry. To achieve a more consistent comparison between the experiment and simulations, new important extensions to the calculations have been recently introduced:

1. The implementation of anharmonic VCD rotational strengths going beyond the double harmonic approximation have been reported [111–113]. This allows the rotational strengths to be calculated not only for fundamental transitions but also for overtones and combination bands.
2. The simulations of VCD and ROA spectra for large molecules, especially for biomolecules, provide a wealth of data. To simplify the interpretation of the calculated vibrational spectra of such molecules and to understand the structure-spectra relationship, a methodology for extracting the so-called localised vibrations has been recently developed [114]. These localised modes are more appropriate than normal modes for the analysis of spectra of proteins and polypeptides.
3. The deepened description of the solvent effects can be found in [115–120]. This important issue is discussesd in more details below.



Scheme 3 The general Scheme of calculations of chiroptical VCD and ROA spectra

5.4.3 Theoretical Modelling of a Molecule in Solution in the Context of Optical Spectroscopy

Most studies on amino acids have been performed in liquid phase, as it is a natural environment for bioprocesses. In liquid phase, the problem of the theoretical modeling acquires an additional dimension: the environment. The majority of the VCD and ROA measurements are taken in solution (in the case of biological species, in aqueous solution), while few of them are provided in m atrix-isolation or gas-phase environments. Until now, there have been no experimental gas-phase or matrix-isolation ROA spectra of amino acids. Therefore, the simulation of solvent effects is of fundamental importance for a fair comparison between the experimental and theoretical data.

Moreover, the conformational changes in the structure as a function of solvent polarity and pH should be followed. A number of studies have shown that the effect of the solvent significantly changes the structures, energies and spectra of a system [65]. Some structures are not stable in the gas phase and gain stability in solvents.

The term "solvation" denotes that the solute molecules M are dissolved in the liquid phase, the "solvent" S. The solvent is considered an assembly of molecules, treated as a continuous, homogeneous, isotropic medium, held together by non-covalent interactions and characterised by specific macroscopic properties (density, dielectric constant, refractive index, etc.). Here, a brief overview of the computational methods that are utilised to study solvation phenomena in VCD and ROA spectroscopy is given. We limit the review to first-principle *ab initio* calculations, and concentrate on the most widespread solvation models.

There are three main groups of methods for evaluating the effects of the surrounding solvent on spectroscopic parameters:

1. The supermolecular approach (SM), i.e. the explicit model, where both the solute molecule and some neighbouring solvent molecules are explicitly included in the quantum mechanics (QM) calculations. The explicit solvent model is especially used to model intermolecular interactions, among them, the formation of a hydrogen-bonded clusters surrounding the solute molecule.
2. The continuum models, i.e., implicit models, are a family of dielectric polarised continuum models (PCM), in which the solvent is modelled as a macroscopic continuum dielectric medium (assumed to be homogeneous and isotropic) characterised by a scalar dielectric constant. The solute, placed in a cavity in a dielectric medium, is described at the QM level, while the solute-solvent interaction is described as the mutual polarisation of solute and solvent.
3. Combined molecular dynamics-quantum calculations, MD/QM, is called the hybrid model. To evaluate the effect of solvation, a classical (MD) or quantum molecular dynamics simulation of the system is performed, and next a number of configurations of the solute molecule plus some solvent molecules, up to a given cut-off distance from the solute, is extracted from the trajectory file. It is then possible to perform a supermolecular calculation (QM) on the cluster and to average the results over a given number of configurations to obtain the spectroscopic parameters in liquid phase. In addition, the second solvation shell could be described by one of the family of PCM models.

If a solute does not form hydrogen bonds with a solvent, the calculations for solution can be performed using a one of reliable PCM approaches. On the other hand, if a solvent molecule forms hydrogen bonds with a solute molecule, the use of PCM solvation models is rather inappropriate. We have restricted our review mainly to the methods of groups (II) and (III), i.e. continuum models and hybrid methods, treating the explicit method only as a reference frame for them. For a detailed description of modelling the solvent effect, we refer the Reader to the other reviews and books [54, 115–120].

5.4.3.1 Supermolecular Approach

To model the solvent effect, one explicitly treats the solute and solvent molecules by quantum chemical methods. Usually only the first solvation shell is considered; therefore, this approach describes only short-range solvent effects. The geometries of the solute–solvent cluster are usually found by means of a geometry optimisation procedure. Optimisation of the global and local minima is confirmed by ascertaining that the harmonic frequencies are real. The conformer population of all minima found is calculated by the use of the Boltzmann distribution under normal conditions and Gibbs free energy differences (ΔG). The interaction energy is corrected for the Basis Set Superposition Error (BSSE) by the counterpoise method [121]. The energies needed to deform the monomer from the optimised monomer geometries to those in the cluster geometry can be estimated by a seven-point correction method [122, 123]. Long-range solvent effects can be accounted for by merging the supermolecular approach and one of the PCM models by placing the solute-solvent cluster in the cavity of a dielectric medium. With an increasing number of solvent molecules, the level of theory is necessarily reduced and nowadays the DFT method is the only choice.

5.4.3.2 Polarised Continuum Models

In this approach, the solvent S is mimicked by a continuous dielectric medium. The solute molecule M (or a group of molecules) is represented by a charge distribution, which occupies a given volume, called the molecular cavity. The solute molecule polarises the solvent, which, as a response, induces an electric field called the "reaction field" (RF), interacting with the solute. The combination of the solute charge, or the field it produces, and the solvent reaction field can be analytically treated to give the interaction potential. Only the effect of the long-range electrostatic interaction of the solute with the solvent reaction field is included in this model.

This general picture can be translated into a quantum mechanical language. The solute is represented by its wavefunction. The interaction with the solvent is introduced in the Hamiltonian through a perturbation, the reaction potential V_R. Including this in the solute Hamiltonian leads to a new, changed solute electronic structure charge distribution. The Self-Consistent Reaction Field method (SCRF) is obtained through an iterative procedure to self-consistency.

Implicit Polarised Continuum Model (PCM) methods based on solvent reaction field philosophy differ mainly in: (1) the cavity shape and (2) the way in which the

charge interaction with the medium is calculated. The cavity is differently defined in the various versions of models (a sphere, an ellipsoid, or a more complicated shape following the surface of a molecule). The cavity should not contain solvent molecules but rather contain, within its boundaries, the solute charge distribution. It should be noted that a drawback of the model is that the results are quite sensitive to the cavity radius. The solvent reaction potential can be partitioned into several contributions of different physical origin, related to electrostatic, repulsive, induction, and dispersion interactions between the solute and solvent.

The most popular of these choices are:

1. PCM methods developed primarily by the Pisa group of Tomasi and co-workers [115, 117, 124], in which only the electrostatic and induction terms are explicitly considered as the interaction potential V_R. The cavity is formed from spheres centred on nuclei in the solute molecule. The electrostatic part of the solvent-solute interaction is represented by the charge density spread on the cavity surface (Apparent Surface Charges, ASC). In the original method, called dielectric D-PCM, the magnitude of the point charges was determined on the basis of the dielectric constant of the solvent.
   Now, the Integral Equation Formalism (the IEF-PCM method), by Cancès, Mennucci and Tomasi is the most popular PCM version [125]. It employs a molecule-shaped cavity composed of spheres centred on the nuclei, while the reaction-field is modelled by placing charges on the cavity surface. To characterise the intermolecular interactions, it is necessary to take into account the non-electrostatic terms. Amovilli and Mennucci have described an approach where the repulsion and dispersion terms are computed self-consistently as part of the reaction field operator [126]. The IEF-PCM method performs better than the reaction-field one, because a more realistic cavity shape is taken into account.
2. The second approach is conductor C-PCM of Barone and Cossi [127], in which the surrounding medium is modelled as a conductor instead of a dielectric medium.
3. The COnductorlike Screening MOdel (COSMO), a modification of the C-PCM method proposed by Klamt and coworkers [128–130], is also extensively used. As suggested by the name, in the COSMO model the solvent is modelled as a conductor while the cavity shape is the same as in the IEF-PCM method.
4. Several other PCMs, i.e., the Multipole Expansion Method (MPE) by the Nancy group of Rivail and Rinaldi [118, 131] and by Mikkelsen and co-workers [132, 133] and the Generalised Born Approximation (GBA) by the Minneapolis group of Cramer and Truhlar [134–136], have also been introduced.

Over the years, a number of PCM models have been used for the calculations and interpretations of the VCD and ROA spectra, differing in the shape of the cavity and in the treatment of the solute-solvent interaction. The majority of PCM calculations of solvent effects on VCD and ROA spectra employ either the IEF-PCM or COSMO approaches. Nevertheless, when the specific solvent-solute interactions become important, the continuum methods are no longer appropriate. In such cases, the explicit supermolecular method in conjunction with the implicit PCM approach should be used.

5.4.3.3 Combined Molecular Dynamics-Quantum Approach (MD/QM)

For systems containing multiple conformations with unknown equilibrium structures and/or for systems with hydrogen-bonded structures between solute and solvent molecules, the hybrid model is needed. There are three ways to use these methods:

1. The simplest method is the supermolecular calculations for solvent molecules from the first solvation shell combined with PCM model for the remaining part of the solvent. It is frequently referred as hybrid method.
2. The majority of classical molecular dynamics (MD) or Carr-Parrinello-Molecular Dynamics (CP-MD) for modelling chiroptical properties in solution employ this approach to generate solute-solvent clusters for supermolecular calculations. Here, the first step is to find the number of solvent molecules hydrogen-bonded to the solute molecule. This can be determined from the analysis of atom-atom radial distribution functions (RDF) or from the generation of solute-solvent clusters as MD "snapshots" [66, 137]. The second step is the conformational search for the cluster and the simulation of a spectrum. Moreover, explicit supermolecular microsolvation calculations can be combined with the implicit PCM method.
3. The most satisfactory and more rigorous approach to account for solvation is to perform dynamic simulations, i.e., by employing MD or CP-MD directly and by simulating the spectra at each time step of the MD trajectories ("snapshots" corresponding to, for instance, 10 ns of simulation). This accounts for the inhomogeneous band broadening and Boltzmann averaging of many conformations provided by the conformer distribution [137–141].

At this point, let us mention an important problem. When the geometry of the solute-solvent complexes is optimised, the transformation of the Hessian to normal coordinates in the frame of the harmonic approximation is correct and the energy gradient with respect to geometric coordinates is zero. However, using "snapshots" leads to a number of imaginary frequencies and thus introduces errors and artifacts in the simulated spectra. A solution to this problem was proposed by Bouř and Keiderling [142] using the scheme of separating the low- and high-frequency regions. At present, such simulations are still very demanding and the investigation of solvent effects on chiroptical properties, using this technique, is limited.

Concluding this part of the review, we will restrict our discussion to the implicit and explicit solvent methods that actually are used to model solute-solvent interactions in VCD and/or ROA spectroscopy for model systems. There are examples showing that PCM models, which despite predicted reasonable energies and geometries [127], do not sufficiently reproduce the vibrational frequencies [143]. Nowadays, most of the VCD and ROA spectra calculations are carried out for biomolecules. Modelling of the hydrogen-bond interactions between solute and solvent molecules requires explicit solvent molecules included in the supermolecular scheme to model the first hydration shell, and then use the PCM approach for the remaining part of the solvent. A good example of the hybrid model can be found

in a study by Jalkanen et al. on the vibrational spectra of the alanine dipeptide [144] and in the investigations of Mukhopadhyay et al. of ROA spectra in aqueous solution using Monte Carlo simulations for water surroundings [145, 146]. As concluded by Jalkanen et al. [144], the hybrid model is the proper approach for further studies of VCD and ROA spectra of amino acids in water. However, very recently, the most promising, yet time-consuming, multiscale approach has been to merge dynamic methods with supermolecular methods and with the PCM model [137–141]. The algorithm contains a few steps: (1) a quantum-chemical exploration of the conformation landscape of the solute molecule (a scan of the potential energy surface, PES) with the PCM model; (2) reoptimisation of the geometry of the complex with a few solvent (water) molecules by quantum methods, or, by a more rigorous approach, finding the geometry of the complexes using the MD or CP-MD methods; (3) calculation of the spectra for the individual complexes and averaging the final spectrum based on the Boltzmann populations of the complexes, or, by a more rigorous approach, calculation of the spectra for hundreds/thousands of MD complexes extracted from snapshots. The final spectra represents the average data. The above mentioned multiscale approach opens up a new generation of methods that can be used for simulations of VCD and ROA spectra, and show better agreement between calculations and experiment than the previously used hybrid methods. The increased computational power of modern parallel computers will make this technique more and more popular.

5.5 Vibrational Chiroptical Spectra of α-Amino Acids in Solution

The number of studies devoted to vibrational circular dichroism or Raman optical activity spectra of amino acids has remained limited and the studies have not been synthesised for all amino acids in a coherent manner. However, a few studies have been dedicated to the spectra of more than two amino acids at once. We begin this section first by referring to the findings of these summarising papers. Therefore, this subchapter is divided into two parts: the first referring to studies on several amino acids and the second to investigations focusing on one or two amino acids.

5.5.1 Vibrational Chiroptical Studies on Common Amino Acid Characteristics

The experimental VCD studies comparing results obtained for several amino acids started in 1983 by the observation that the C*H VCD band of L-amino acids in aqueous solution exhibits a strong positive bias [147]. In that study, all natural α-amino acids (except aspartic and glutamic acids, tyrosine, and tryptophan, but additionally allo-threonine, 4-hydroxy proline, and allo-4-hydroxy proline) and penicillamine

were registered. Since then, the Nafie group have studied this problem in more and more detail [148–150]. In 1985, they explained the positive sign of the C-*H band by the formation of intramolecular hydrogen bonds between the ND_3^+ and CO_2^- groups [148]. In 1988, they observed a characteristic negative/positive band VCD pattern in the 1200–1400 cm^{-1} range of the VCD spectra of L-alanine, L-proline, and selected dipeptides dissolved in water [149]. The negative and positive band were assigned to two orthogonal bending modes of the C*H(CO_2^-) group of L-amino acids and peptides. In 1989, the VCD spectra of L-alanine and its deuterated isotopomers, L-serine, L-cysteine, and β-chloro-L-alanine were again analysed in the C*H stretching vibrations region [150]. The VCD spectra were examined as a function of pH and it was found that at neutral and high pH, the spectra exhibited a large positive VCD intensity bias associated with the C*H stretching mode, whereas at low pH, the bias was absent and only very weak VCD signals were detected. At that time, the interpretation of the VCD intensity could only be supported by a general consideration of amino acid conformation but not confirmed by quantum chemical calculations.

VCD analysis of amino acids in aqueous solution in the mid-IR region is difficult to undertake since (1) high concentrations and short path lengths should be used because of the strong absorption of water and (2) amino acid solubilities in water at neutral pH are usually not sufficient for VCD measurements. A small experimental breakthrough in the registration of the VCD spectra of amino acids in water was achieved in 2005 by the innovation of a film technique by Shanmugam and Polavarapu [151]. The technique consists of the formation of aqueous films of water-soluble α-cyclodextrin as a matrix. In this way, the strong interference of water absorption is eliminated, enabling the VCD technique to be applied to biologically important compounds which are poorly soluble in water. The innovation lies in the fact that, even though moderately absorbing α-cyclodextrin provides a chiral environment, its interaction with amino acids is weak [152] and the VCD spectra of amino acids are practically unperturbed. In 2006, the VCD spectra of L-alanine, L-proline, L-methionine, L-histidine, L-phenylalanine, and L-tryptophan in the form of aqueous films with α-cyclodextrin were registered by Zhang and Polavarapu [38, 153]. A wide spectral region of the VCD spectra of L-methionine and L-histidine, as well as poorly soluble in water L-phenylalanine and L-tryptophan, were registered for the first time [151]. For moderately soluble L-alanine, L-proline, L-methionine, and L-histidine, the VCD spectra of films with α-cyclodextrin were compared with those obtained in the solution state and good agreement was found. For L-tyrosine, neither absorption nor VCD spectra could be obtained in the solution state because of extremely low solubility.

In 2006 Jalkanen et al. published complex, combined computational research on IR, VCD, Raman, ROA, and SERS spectroscopy of L-alanine, L-tryptophan, L-histidine, some model peptides, and a variety of small molecules of relevance to protein systems [154]. A variety of computational methods including molecular mechanics, SCC-DFTB (self-consistent charge density functional tight binding) [155], RHF, MP2, and DFT methods were combined with continuous and discrete solvation by several water molecules. The discrete hydration model (four water molecules

and the Onsager continuum model for bulk water) enabled the authors to find a computationally stable L-alanine zwitterion due to the network of water molecules stabilising and bridging the two charged groups. Without the explicit water molecules, one finds only the neutral conformers to be stable [156]. Also, for L-histidine and L-tryptophan, four water molecules were necessary to stabilise their zwitterions and to predict their VCD and ROA spectra. However, no other generalisations about the VCD and ROA spectra amino acids other than those connected to the proper reproduction of their zwitterionic forms were proposed. Now, the development of computer technology and quantum chemical software abilities is so fast that, often, already after a few years, the computational results become outdated. Nevertheless, it seems that the conclusions of the Jalkanen et al. investigations [154] still remain valid.

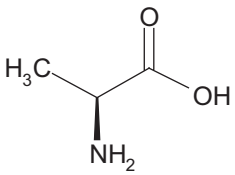
Recently, the VCD and IR spectra of the 20 coded L-amino acids were calculated at the B3LYP/cc-pVTZ level of theory [157]. Despite the significant non-rigidity of amino acid molecules [2], the calculations were based on single structures available from a database [158]. A fair agreement between the calculated and experimental spectra was declared; however, for the IR spectra, the generalising conclusions were similar to those that could be drawn by inspection of the classical correlation tables [159, 160]. For the VCD spectra, for which there is still a lack of generalised rules about changes in VCD intensity upon changing different factors (solvent, H-bonding, conformation, etc.), it was found that the ρ(OH) rocking and ν(C=O) stretching modes in most amino acids exhibit the largest VCD rotational strengths. It would be worth verifying this important conclusion at a sufficiently high level of theory because the correct reproduction of the amino acid spectra collected even in slightly interacting medium such as low-temperature matrices requires a level of theory including the electron correlation effects and a large basis set [161–164].

The ECD and VCD spectra of model compounds mimicking aromatic amino acid residues in a protein (L-phenylalanine, L-tyrosine, L-histidine, and L-tryptophan with the -COOH and -CNH_2 groups replaced by -$CONHCH_3$ and -$CNHCOCH_3$) were studied by Tanaka et al. [165]. Quite a variety of conformations of each residue were calculated at the TD-DFT and DFT levels, and the spectra were superposed to show how the conformations influence the spectral picture. The VCD spectra appeared to be more sensitive to side chain conformations than the ECD spectra, and it was indicated that using DFT calculations, one can elucidate the main and side chain conformations of the aromatic residues in proteins [165].

Other studies on the VCD and ROA spectra of amino acids have been very unsystematic, so far. There are numerous thorough studies on L-alanine spectra, both experimental and theoretical, and several papers have been devoted to L-proline and some of its derivatives, while only a few papers describe the spectra of cysteine and serine. The VCD and ROA spectra of the other amino acids have been analysed sporadically if at all. Below, we systematically discuss the experimental and theoretical VCD and ROA studies on amino acids starting from spectroscopic investigations of L-alanine and L-proline.

5.5.2 Vibrational Chiroptical Studies on Particular Amino Acids

5.5.2.1 L-alanine



Scheme 4 2-aminopropanoic acid

VCD: Experiment The most recent experimental and computational study on the VCD spectra of alanine in heavy water (1250–1800 cm^{-1}), known to us, was the subject of BSc thesis at the University of Warsaw [166]. However, in 1977, the first reliable registration of the VCD spectra of L- and D-alanine-d_3 in heavy water in the CH stretching vibrations region (3100–2850 cm^{-1}) was a very challenging task [167]. Since then, alanine VCD spectra have been repetitively measured [147–149, 168–176] and calculated [144, 171, 172, 175, 177–183].

The most complete early measurements of the VCD spectra of L-alanine and L-alanine-N-d_3 in H_2O and D_2O were done by Diem for the 1700–900 cm^{-1} region [174]. Beyond the C*H bending vibrations studied before [149], weak negative and weak positive vibrations at 1623 cm^{-1} and 1418 cm^{-1} were observed and assigned to asymmetric and symmetric stretching vibrations of the CO_2^- group, respectively. The broad positive peak at 1221 and the broad negative shoulder at 1139 cm^{-1}, which disappeared upon deuteration, were assigned to the rocking mode of the NH_3^+ group. Finally, a negative band at 1117 cm^{-1} was assigned to asymmetric stretching vibrations of NCC heavy atoms in the $CH_3CHNH_3^+$ moiety.

In 2003, Ellzy et al. collected up to 50,000 scans of the VCD spectra of L- and D-alanine pressed as a KBr pellet and in water (c = 1 M, path length of 0.006 mm, BaF_2 windows) [175]. They concluded that the experimental VCD spectrum of L-alanine taken from the KBr disc compared favourably with the B3LYP/6-31G* spectrum [175].

Wang and Polavarapu were probably the first to introduce the two-dimensional VCD (2D VCD) technique [184]. They analysed the details of the interaction between avidin, a glycoprotein isolated from egg-white, with biotin (vitamin H) [184]. Even though the 2D VCD technique is still rarely used [185, 186] in our opinion it certainly will be, along with two-dimensional Raman optical activity spectroscopy (2D ROA) [187, 188], an indispensable method for analysing the conformational complexity of biosystems by vibrational spectroscopy methods. The Ma et al. study in 2006 is an extraordinary example of the potency of 2D VCD spectroscopy in examining the structure of L-alanine in water as a function of pH [176]. In that work, the 2D correlation spectroscopy (2D-COS) was coupled to VCD and IR spectra in multiple spectral regions. In addition, the excellent one-dimensional VCD and IR

spectra of L-alanine in D_2O were recorded in the mid-IR (800–2000 cm^{-1}) region, the hydrogen-stretching region (2000–4000 cm^{-1}), and, for the first time, the near-IR region (4000–6000 cm^{-1}) as a function of pD (Fig. 5.1). A variety of different 2D-COS plots were demonstrated based on IR and VCD spectral data, as well as the combination of both IR and VCD spectra. Analysis of homo- and hetero spectral regions of VCD–VCD and IR–IR 2D-COS plots revealed the correlation of mutual changes of methine and methyl bending, methine and methyl stretching and carboxylate stretching regions with changes in the acidity and basicity of the solution. The 2D maps evidently enhanced the resolution of overlapping bands and assignments of combination bands especially in the NIR VCD region.

In 2006, the VCD bands of L-alanine, described before as weak [174], were called strong [176]. The characteristic negative/positive intensity pattern of two orthogonal C*H bending modes at 1337 cm^{-1} and 1291 cm^{-1}, respectively, as well as the weak negative CH_3 umbrella mode at 1377 cm^{-1}, exhibit most of the spectral changes at low pD between 3.0 and 1.5, whereas in the pD range from 3.0 to 13.0, the changes occur smoothly (Fig. 5.1a, b). For higher frequency regions, only the spectra for high pD were measured (Fig. 5.1c, d)

In comparison to IR, the corresponding VCD intensities are four to five orders of magnitude smaller, and, as a consequence, the VCD–VCD 2D-COS plots are eight to nine orders of magnitude smaller than the corresponding IR–IR plots. Moreover, the synchronous 2D-COS plots are usually at least an order of magnitude larger than the associated asynchronous plots. Here, we mention only some of the 2D maps generated in Ref. [176]. The negative cross-peaks at 1291 and 1337 cm^{-1} of the synchronous VCD–VCD 2D-COS plot (Fig. 5.2a) represent the pD-induced correlation of the two bending motions of the C*H group. They indicate that the two correlated bands changed intensity in the opposite direction. This means that they both decreased in intensity. The corresponding asynchronous cross-peaks at 1291 and 1337 cm^{-1} were positive (yellow, Fig. 5.2b), meaning that, over the range of the pD perturbation, the lower frequency C*H bending motion changed earlier than the higher frequency motion. High pD changed the charge on the amino group. Thus, C*H movement towards the NH_3^+ group was more sensitive to pD changes than those directed towards the carboxylate group. This indicates that the C*H bending positioned at 1291 cm^{-1} corresponds to modes directed towards the NH_3^+ group, whereas the mode at 1337 cm^{-1} can be assigned to motions towards the CO_2^- group. Similarly, the weak positive synchronous and asynchronous cross-peaks (1377 and 1337 cm^{-1}) indicate that the umbrella $\delta^s(CH_3)$ mode (1377 cm^{-1}) is more sensitive to high pD than the higher frequency bending δ(C*H) mode. The synchronous and asynchronous VCD–IR 2D-COS of L-alanine bands at 1291 and 1337 cm^{-1} (Fig. 5.3) lead to the same conclusion as those derived from the VCD–VCD 2D spectra i.e., that the methine group is more sensitive to high pD than the methyl group to pD changes.

In comparison to conventional 1D VCD spectra, a significant resolution enhancement was obtained for the 2D NIR VCD spectra (Fig. 5.4). Thanks to this, by comparing the signs of synchronous and asynchronous 2D VCD peaks, the authors could obtain the following sequences of apparent change as a result of increasing

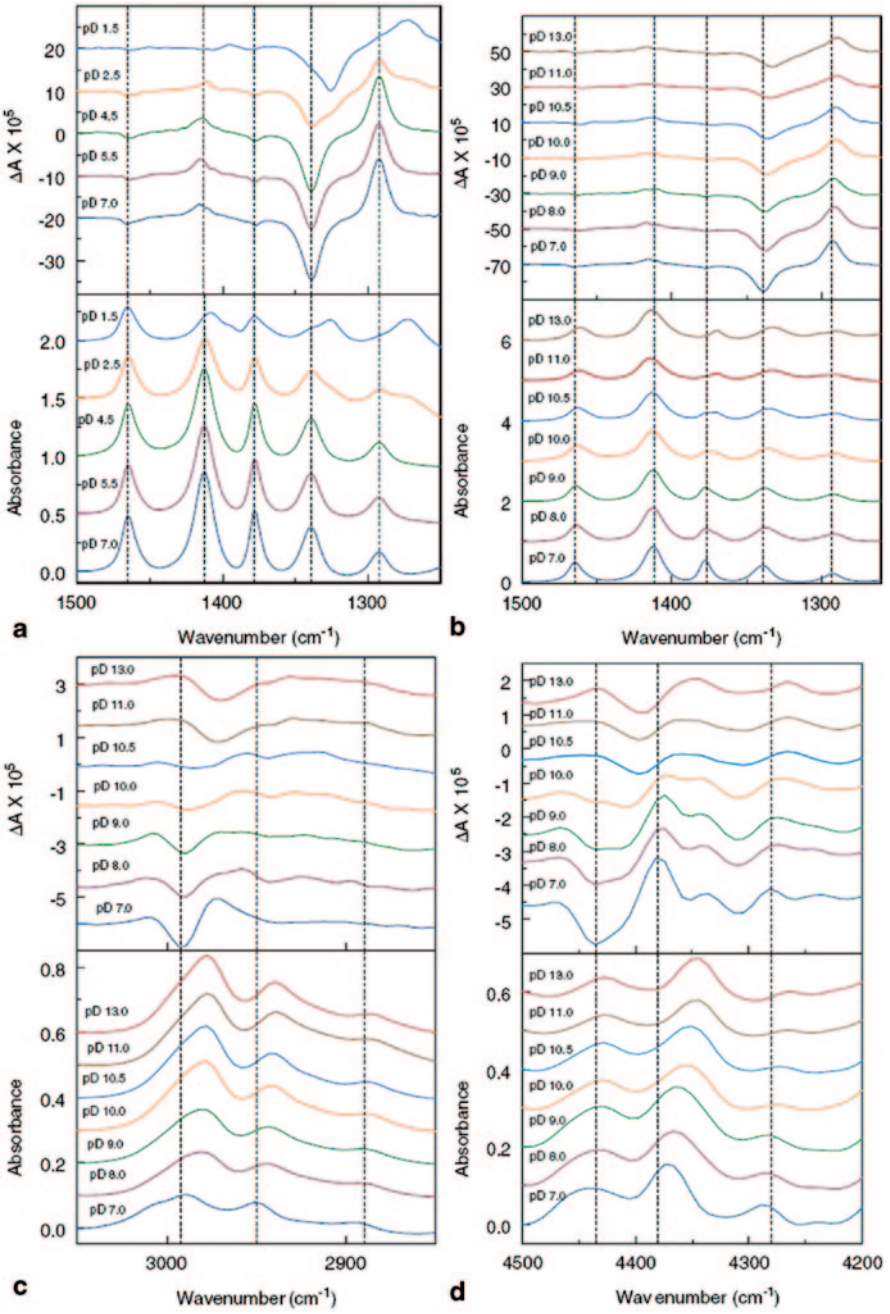


Fig. 5.1 A stack of VCD spectra (*upper traces*) and absorption spectra (*lower traces*) of L-alanine in the: CH-bending region (1500–1250 cm^{-1}, 30 μm path length) under low (a) and high (b) pD conditions, (c) in the CH-stretching region (2800–3100 cm^{-1}, 50 μm path length) and (d) in the NIR region (4200–4500 cm^{-1}, 2 mm path length). Experimental details: 1.7 M in D_2O solution and 4 cm^{-1} resolution. (Reproduced from Ref. [176] with kind permission of Elsevier)

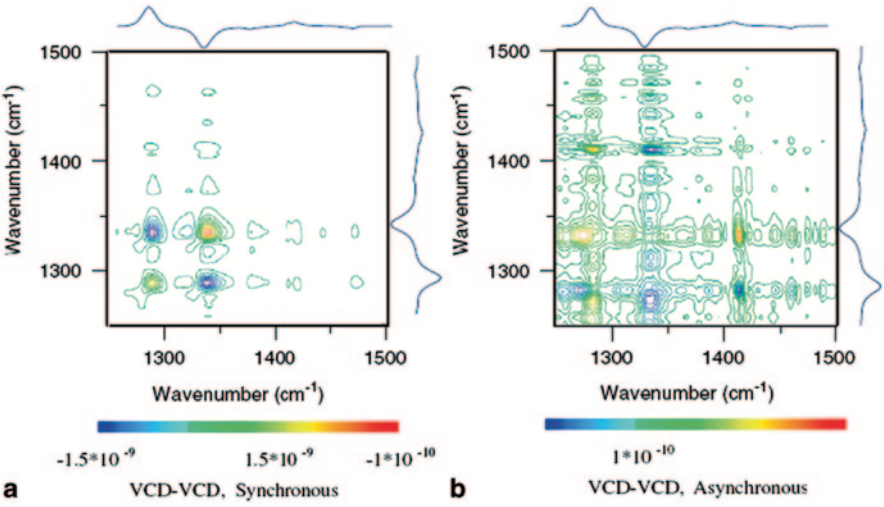


Fig. 5.2 (**a**) Synchronous and (**b**) asynchronous 2D VCD COS in the 1250–1500 cm^{-1} under high pH (from 7 to 13) conditions. 1D spectra at pD = 7 are plotted on the *top* and *side*. A colour bar is placed on the *bottom* to show the correlation intensities. (Reproduced from Ref. [176] with kind permission of Elsevier)

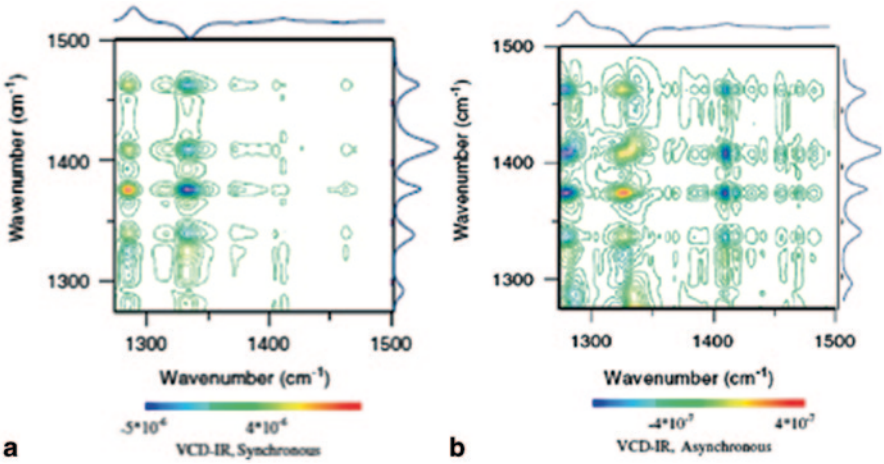


Fig. 5.3 (**a**) Synchronous and (**b**) asynchronous 2D VCD-IR COS in the 1250–1500 cm^{-1} region under high pH (from 7 to 13) conditions. 1D spectra at pD = 7 are plotted on the *top* and *side*. A colour bar is placed on the *bottom* to show the correlation intensities. (Reproduced from Ref. [176] with kind permission of Elsevier)

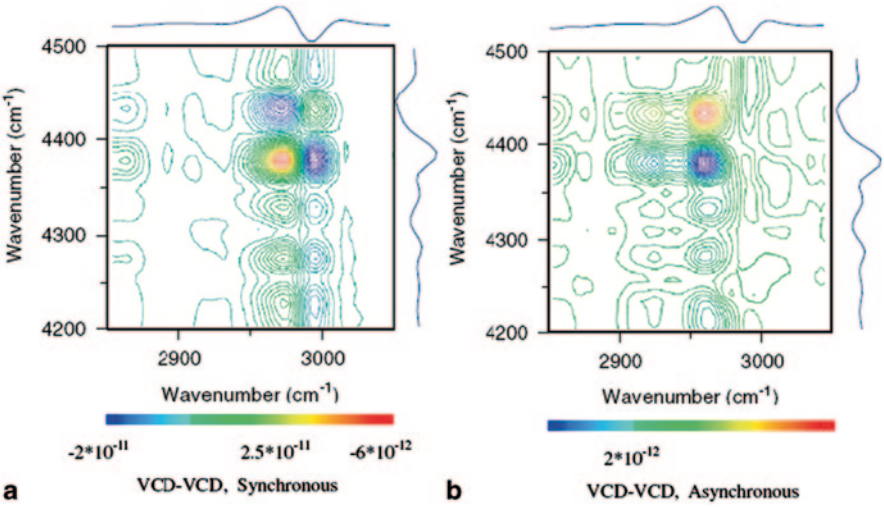


Fig. 5.4 (**a**) Synchronous and (**b**) asynchronous 2D VCD COS between the CH-stretching and NIR regions (2800–3100 cm^{-1}) to (4200–4500 cm^{-1}). (Reproduced from Ref. [176] with kind permission of Elsevier)

the pD: 4266 < 4435 < 4381 < 4312 cm^{-1}. The hetero-region of the VCD–VCD correlation plots (region of different frequency ranges at the two coordinates) showed a strong positive synchronous cross-correlation intensity at 3011 and 4381 cm^{-1} in the VCD–VCD correlation plot, which indicated that the combination band at 4381 cm^{-1} included the high frequency member of antisymmetric $\nu^a(CH_3)$ mode (3015 cm^{-1}). As a result, it was possible to assign the band at 4381 cm^{-1} to the combination of CH_3 deformation (1377 cm^{-1}) and methyl stretching (3015 cm^{-1}) modes.

Based on changes observed in the 1D and 2D VCD and IR spectra of L-alanine in heavy water solution perturbed by pD changes from 1.5 to 13, the authors concluded that all changes in the spectra as pD changed were the result of a change in the relative population of the zwitterion (neutral solution) and the anion (high pD) or the cation (low pD). For the measured VCD and IR spectra, no evidence of dimeric species formation was uncovered. The pK_a values for alanine in these two pD regions are approximately 2.7 (low pD) and 10.3 (high pD). The 2D analysis of cross-correlation peaks of synchronous and asynchronous spectra provides information on whether one spectral region or another changes at different rates as a function of pD changes, and it only appears as though the first band changes in intensity ahead of the second band. The spectra presented the sensitivity of 2D-COS spectroscopy across a wide range of spectral types and frequency ranges to analyse subtle changes in the spectra of a molecule undergoing a transition from one species to another. Features are resolved in the 2D spectra that are not at all apparent in the simple 1D spectra.

VCD spectra: Calculations Probably the first calculations of the VCD spectra of three isotopomers of alanine were done by Diem et al. and Lal et al. in 1982 [171, 172] at the semiempirical CNDO/2 (Complete Neglect of Differential Overlap) level [189]. It is not strange that at that time "a very close relative agreement" between semi-empirical and experimental spectra in the CH stretching vibrations region was stated and the discrepancy between theoretical and experimental spectra was attributed to the presence of water influencing the spectra. It is noteworthy that after 30 years of theoretical calculations of alanine vibrational optical activity spectra, a proper inclusion of the influence of water on the VOA spectra is still the most important issue [141, 154, 190, 191]. The first *ab initio* HF/SCRF/6-31G* calculations of the VCD (and ROA) spectra of alanine were presented by Nafie et al. in 1994 [179]. The influence of surrounding water was simulated using the self-consistent reaction field (SCRF) which modelled the water environment by setting the relative permittivity to 80 outside the cavity with a radius of 3.5 Å in which the alanine molecule was closed. The calculated spectra were compared with those measured by Diem [174]. It is not surprising that, at that level of theory, some bands were misordered and produced assignment problems. Nevertheless, according to the prevailing standards of the 1990s, reproduction of the experimental spectra was again evaluated as being very good [179]. The strength of the Polarised Continuum Models in theoretical studies of the solvent effect on the VCD spectra of alanine was presented by Gontrani, Mennucci, and Tomasi in 2000 [180]. They demonstrated that PCM combined with the B3LYP/6-31G* level of theory was able to reproduce both the stability of zwitterionic forms of alanine in water solution and the VCD spectra, except for the modes in which alanine H-atoms form hydrogen bonds with water. The authors suggested that a supermolecular approach such as that already used a few years earlier by Tajkhorshid et al. [181] could improve the predictions of the VCD intensity of those modes.

The supramolecular model of the solvent effect on the VCD spectra of L-alanine was repetitively applied by Tajkhorshid et al., [181] Frimand et al., [182] and Jalkanen et al. [144, 154] at gradually increasing levels of theory or complexity of the studied supermolecular system. In the first of this series of papers [181], it was found that the Onsager solvent model cannot stabilise the alanine zwitterion against conversion into the neutral tautomer, whereas a discrete set of four water molecules stabilises zwitterionic alanine in the B3LYP/6-31G* calculations. To calculate the VCD spectra, the distributed origin gauge atomic axial tensors and the electric dipole electric-dipole polarisability derivatives were calculated at the RHF/6-31G** level of theory for the B3LYP calculated geometries. The as-calculated VCD spectra appeared to be very sensitive to the relative arrangement of alanine and the neighbouring water molecules. In the next study [182], also performed at the RHF and DFT levels combined with the 6-31G* basis set, the AATs were calculated with the B3LYP functional. The alanine zwitterion was considered in three ways: (1) as surrounded by explicitly four or nine water molecules only, (2) embedded in a self-consistent reaction field expressed by the Onsager model, and (3) surrounded by explicit water molecules embedded in a self-consistent reaction field. It was found that most bands of the VCD and IR spectra in the range of 1650–840 cm^{-1},

corresponding to symmetric, anti-symmetric and rocking modes of CH moieties, were not significantly affected by water, while the C=O band located in the harmonic model above 1650 cm^{-1} was sensitive to surrounding water. The authors concluded that only treating the solvent effects by adding both explicit water molecules and surrounding them by a dielectric media gave a really fair comparison to experimental VCD spectra [182]. The following paper was devoted to calculations of the VCD spectra of the alanine zwitterion deuterated at C-atoms and used essentially the same methodology as before [183] (Fig. 5.5).

In the most recent paper by Jalkanen et al. [144] devoted to the theoretical VCD spectra of L-alanine, DFT Born–Oppenheimer molecular dynamics simulations were performed to determine the locations of the water molecules in the first solvation shell responsible for stabilising the zwitterionic structure [190]. It was found that as many as 20 water molecules were necessary to fully encapsulate zwitterionic alanine inside a droplet of water, while 11 water molecules were necessary to encapsulate the polar region of alanine, exposing the hydrophobic methyl group to the droplet surface, where it migrated during the MD simulations. The bulk water was simulated by different models such as Onsager, PCM, and COSMO. The agreement between experimental and calculated spectra (Fig. 5.5) shows that the use of explicit water molecules and continuum solvent treatment is very satisfactory for simulations of L-alanine and, as we shall see further, for other amino acids, even if the basis set used is not large. The series of studies on L-alanine and related biological compounds with Jalkanen as a co-author are summarised in Ref. [192]. The problem of a proper computational representation of hydration of the alanine zwitterion is still a hot issue [144, 190, 191] and it is very likely that the last word has not been said, yet.

ROA spectra: Experiment and Calculations In the case of the ROA spectra of L-alanine, early measurements were accompanied by quantum chemical calculations. The first detection of the Raman optical activity (ROA) spectra of L-alanine in water solutions was performed by Hecht et al. in 1989 [193] and Barron et al. presented the reliable spectra in 1991 [156]. The spectra, recorded in backscattering mode for the 500–1700 cm^{-1} region in neutral, basic and acidic solutions, were found to be relatively insensitive to pH changes (from 1 N HCl to 1 N NaOH). However, a broader range of pH (from 6 N HCl to 6 N NaOH) enabled the authors to observe significant changes in the ROA spectra, to reassign the L-alanine bands and to refine previous conclusions [194]. It was emphasised that the presence of ROA structure in regions, where none is seen in the parent Raman spectra, indicate increased resolution capability. The experimental ROA intensities were compared with dimensionless circular intensity differences [100] calculated for the L-alanine zwitterion at the SCF level combined with the 6-31G and 6-31G* basis sets. A remarkable agreement of the calculated intensities with the experimental parameters in the lower-frequency region was reported.

In 1994, Nafie et al. compared VCD and ROA of L-alanine in water solutions [179]. A good agreement between the SCF/6-31G* calculated spectra combined with the Onsager self-consistent reaction field and the measured ROA as well as VCD

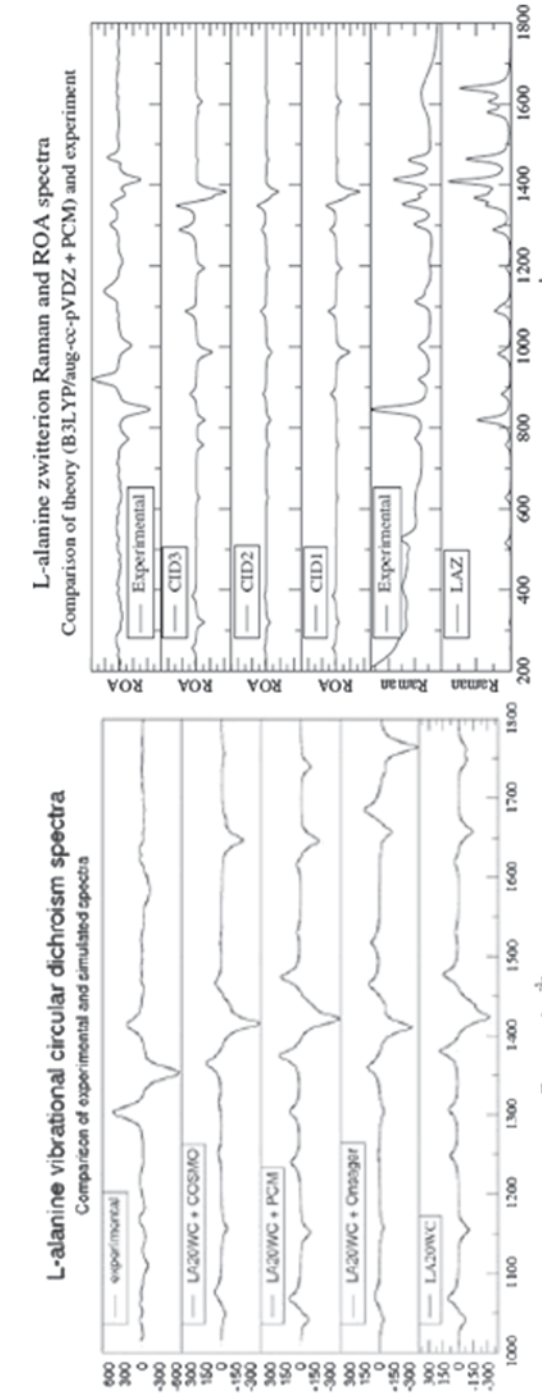


Fig. 5.5 A comparison of the B3LYP/6-31G* calculated VCD (**a**) spectra for L-alanine surrounded by: 20 water molecules (LA20WC), 20 water molecules in the first solvation shell and the Onsager model for the bulk water (LA20WC + Onsager), 20 water molecules in the first solvation shell and the PCM model for the bulk water (LA20WC + PCM) and 20 water molecules in the first solvation shell and the COSMO model for the bulk water (LA20WC + COSMO) and the experimental spectra. (**b**) Comparison of theoretical and experimental Raman optical activity spectra for the L-alanine zwitterion. (Reproduced from Ref. [144] with kind permission of Springer Science + Business Media)

spectra was concluded. The authors emphasised that the ROA and VCD spectra did not resemble one another and are complementary in content. However, it was argued that ROA has greater potential than VCD for the study of biological molecules in aqueous solution because ROA can be measured in all spectral regions for light and heavy water while this is not possible for VCD. For example, the spectral region between 1500 and 1700 cm^{-1} in VCD spectra has to be measured in D_2O because of the strong absorbance of HOH bending vibrations in liquid H_2O. On the other hand, ROA in aqueous solution can be measured without hindrance down to 100 cm^{-1} [179].

Backward scattering in-phase dual circular polarisation (DCP_I) Raman optical activity (ROA) spectra of L-alanine and its -2-d_1 and -3,3,3-d_3 isotopomers were measured in light and heavy water in the 1700–800 cm^{-1} frequency region by Yu et al. in 1995 [195]. The registration of ROA for deuterated isotopomers followed by theoretical interpretation of the spectra enabled the authors to evaluate the contributions from the motion of deuterated C*H and CH_3 groups as well as CO_2^- and NH_3^+ in generating large ROA signals. The spectra were also calculated at the SCF/6-31G* level with the Onsager reaction field used for stabilisation of the zwitterionic forms. Excellent agreement was found between experiment and calculation for both Raman and ROA spectra, except for alanine -2-d_1. The disagreement was assumed to be an effect of too low a level of calculations applied in the study. The experimental and theoretical ROA spectra of a 1.6 M solution of L-alanine in neutral H_2O were also shown to be in fair agreement in the range of 1300–750 cm^{-1} if the SCF/6-31G* method was applied, even without additional SCRF simulations in the water medium [196].

The ROA spectra of L-alanine in aqueous solution were next predicted by Jalkanen et al. using the HF and B3LYP methods and the 6-31G* basis set [197]. However, the electric dipole–magnetic dipole polarisability derivatives (EDMDPD) and electric dipole-electric quadrupole polarisability derivatives (EDEQPD) required to obtain the ROA spectra, at that time, had to be obtained at the RHF/6-31G*/STO-3G level. The L-alanine structures optimised earlier with and without the explicit presence of four water molecules [182] were used to calculate the ROA spectra, whereas the mixed basis set approach was used for those surrounded by nine water molecules; the L-alanine zwitterion was calculated with the 6-31G** basis set while the water molecules were calculated with the STO-3G set [197]. The Onsager model was considered for the studied structures, but the influence of bulk water on EDMDPD and EDEQPD was ignored. In conclusion, the L-alanine zwitterion surrounded by nine water molecules and the bulk water modelled in the frame of the Onsager model appeared to be the best for reproduction of the experimental spectra obtained by Barron et al. [156, 194] and Yu et al. [195]. In the next article, modelling the ROA spectra of the L-alanine zwitterion in water was performed by the supramolecular approach using four water molecules for the first solvation sphere and the PCM continuum model for the bulk water [144]. This study, performed at the B3LYP/aug-cc-pVDZ basis set, led to better compatibility with newly measured experimental ROA spectra (Fig. 5.5) than those previously obtained using the 6-31G* basis set [154].

The role of the rotation of CH_3, NH_3^+, and CO_2^- groups on the ROA spectrum was estimated by Kapitán et al. [198] using the B3LYP/CPCM/6-31++G** level of theory for structural details and the HF/6-31++G** method for the intensity tensors for a zwitterion in a vacuum (Fig. 5.6a). The band shapes were obtained by Boltzmann

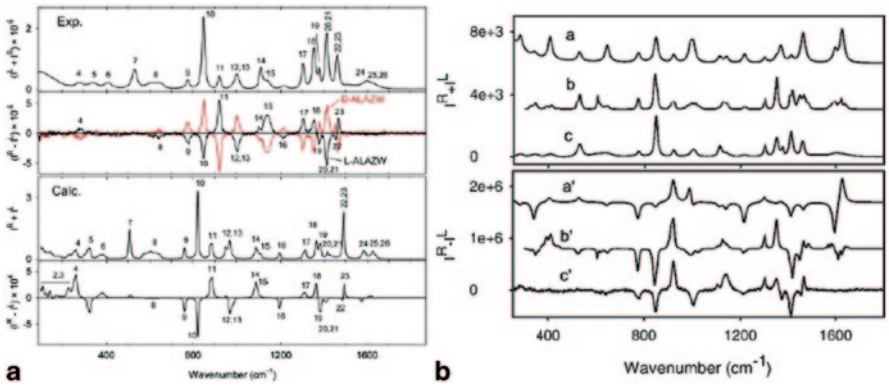


Fig. 5.6 (**a**) Comparison of experimental Raman and ROA ICP spectra of the L-alanine zwitterion in solution with simulations obtained using the Boltzmann average over rotation of the CH_3, NH_3^+, and CO_2^- groups. Experimental ROA spectra of both enantiomers are plotted, while only the L-form was simulated (Reproduced from Ref. [198] with kind permission of American Chemical Society). (**b**) The Raman (a–c) and ROA (a'–c') alanine spectra calculated *ab initio* at the B3LYP/6-311++G**/CPCM level for the lowest-energy conformer (a, a'), MD averaged spectra (b, b'), and the experiment (c, c') (Reproduced from Ref. [138] with kind permission of John Wiley and Sons)

averaging of the bands of individual rotamers. The bands exhibited different sensitivities to group rotations and usually became broader for Raman and weaker for ROA. Very importantly, the ROA signs of about half of the total number of the bands were reversed by the three studied rotations. The disagreement between simulated spectra with the experimental ones in the C-H and N-H bending region (1500–1415 cm^{-1}) was thought to be an effect of anharmonic interactions perturbing these very motions. In the next study by Daněček et al., [199] the role of anharmonic terms in the potential and intensity tensors was systematically explored by vibrational self-consistent field, vibrational configuration interaction (VCI), and degeneracy-corrected perturbation calculations. The harmonic frequencies obtained using 23 DFT functionals and the HF and MP2 methods, the 6-31++G** basis set, and the CPCM model of bulk water were compared to experimental band positions. The harmonic approach appeared satisfactory for vibrations in the 1800–200 cm^{-1} region. The anharmonic corrections included by numerical differentiation did not lead to complete agreement between the experimental and calculated ROA spectra. However, for the C–H stretching and N–H and C–H bending modes, a significant improvement in the Raman and ROA spectral profiles was achieved with use of the VCI correction. A limited Boltzmann averaging for the lowest-frequency modes that could not be included directly in the anharmonic calculus provided realistic inhomogeneous band broadening. The anharmonic parts of the intensity tensors (second dipole and polarisability derivatives) were found to be less important for the entire spectral profile than the force field anharmonicities (third and fourth energy derivatives), except for a few weak combination bands which were dominated by the anharmonic tensor contributions. It was concluded that the assignment of the experimental spectra for the 1550–1205 cm^{-1} range were still not reliable, but this time the authors explained this fact by the high density of C–H bending

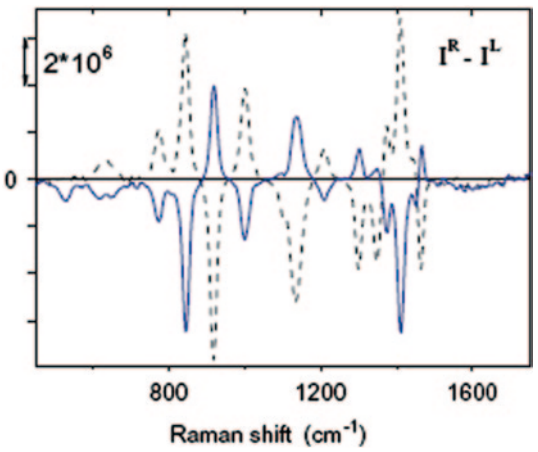


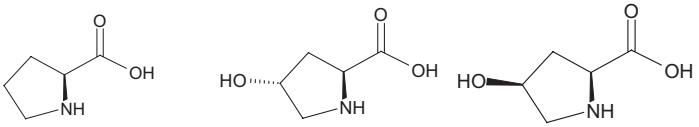
Fig. 5.7 The ROA spectra of zwitterionic L-alanine in water obtained using 532.5 nm excitation. The D-alanine spectrum is shown in dashed lines to show that there are no artifacts. (Reproduced from Ref. [200] with kind permission of Elsevier)

vibrational states and coupling to the motion of the polar groups strongly interacting with the environment [199].

In the most recent paper by Kaminský et al. [138], the methodology of inclusion of multiconformational equilibria in the ROA calculations of the alanine zwitterion was improved (Fig. 5.6b). An additional spectral overlap gradient term was added to the molecular energy during the molecular dynamics run. An alternative algorithm based on the comparison of the averaged spectra with the reference, enabling iterative updates of the conformer probabilities, provided realistic ROA spectra and conformer alanine distributions in a shorter period of time.

Very recently, Fang et al. demonstrated an algorithm deriving the bond polarisabilities from Raman intensities and differential bond polarisabilities from ROA intensities. The two polarisabilities were applied to understand the electric and magnetic coupling in ROA of the model (+)-(R)-methyloxirane and the L-alanine zwitterion in water (Fig. 5.7) [200].

5.5.2.2 L-proline



L-proline
(2*S*)-pyrrolidine-2-carboxylic acid

4-hydroxy-L-proline
(2S,4R)-4-hydroxy proline

allo-4-hydroxy-proline
(2S,4S)-4-hydroxy proline

Scheme 5 L-proline and derivatives

VCD spectra: Experiment and Calculations L-proline is a relatively small, rigid and water-soluble amino acid which, probably, is the main reason why the number of VCD and ROA studies dedicated to it is greater than for the other natural amino acids except for L-alanine. The VCD C*H stretching and bending vibrations bands of L-proline (along with 4-hydroxy and allo-4-hydroxy proline) were measured in water in the 1980s, revealing a pattern characteristic of all L-amino acids [147–149].

Therefore, it is surprising that the first systematic study of the VCD spectra of L-proline in water was a computational study by Cappelli et al. in 2005 [201]. The authors investigated the VCD and IR spectra of L-proline in water using the DFT method, with three water molecules in the first solvation sphere with or without the IEF-PCM version of the polarisable continuum model of bulk water. Four forms of L-proline (neutral, zwitterionic, cationic, and anionic) were considered. The authors found that hydrogen bond effects on spectra can only be partially reproduced by the sole continuum solvent model, whereas the introduction of explicit water molecules shifted several modes towards lower frequencies, changed the intensities and, for some of VCD bands, also changed the sign. Obviously, protonation was also shown to noticeably change the spectra. Interestingly, the environment affected the geometry relatively little, but sometimes dramatically changed the Boltzmann population of the conformers.

In 2006, the VCD spectra of L-proline (and several other natural amino acids) were measured in the form of aqueous films with α-cyclodextrin (α-CD, Fig. 5.8a) [38]. A negative VCD band at 1628 cm^{-1} due to stretching vibrations of the C=O group could be registered more easily than the bands at ca. 1550 cm^{-1}, because the L-proline spectra are weak and the absorbance of α-CD bands in the 1200–1000 cm^{-1} region is strong. Therefore, in this range, the VCD bands of L-proline were not very reliable. The 1550–1200 cm^{-1} region of the VCD bands could be seen in the subtracted spectra in which the appropriate solution of racemate proline in α-CD film was used as a reference background. Between 1400 and 1300 cm^{-1}, a negative-positive doublet, reported earlier [149], was clearly seen in the subtracted VCD spectrum and confirmed in the IR absorption spectra (Fig. 5.8a). It worth adding that in 2004 and 2006, the VCD and IR spectra of N-acetylproline in water were measured and calculated to determine its conformation in water and to verify the MD force fields used for larger amides and polypeptides [202, 203].

Quite recently, Lüdeke et al. recorded the VCD spectra with a tunable external-cavity quantum cascade laser (QCL) [204]. They showed that difficulty in standard VCD spectroscopy measurements due to strong IR absorption of solvents can be overcome by using QCLs, first demonstrated in 1994 [205]. QCLs provide at least 100 times more power in the MIR than ordinary thermal IR light sources and permit the use of IR detectors that do not require N_2 cooling. The wavelength of a QCL can be tuned within a window of ca. 8% around the central wavelength, yielding a ca. 100 to 200 cm^{-1} spectral range lying between 800 and 2500 cm^{-1}. However, the spectral range of 100 cm^{-1} may be sufficient to cover important difference bands such as amide I or amide II bands of peptides. The authors demonstrated the 1320 and 1220 cm^{-1} range of the QCL-VCD spectra of L- and D-proline taken in water with background correction by a racemic mixture. In agreement with previous

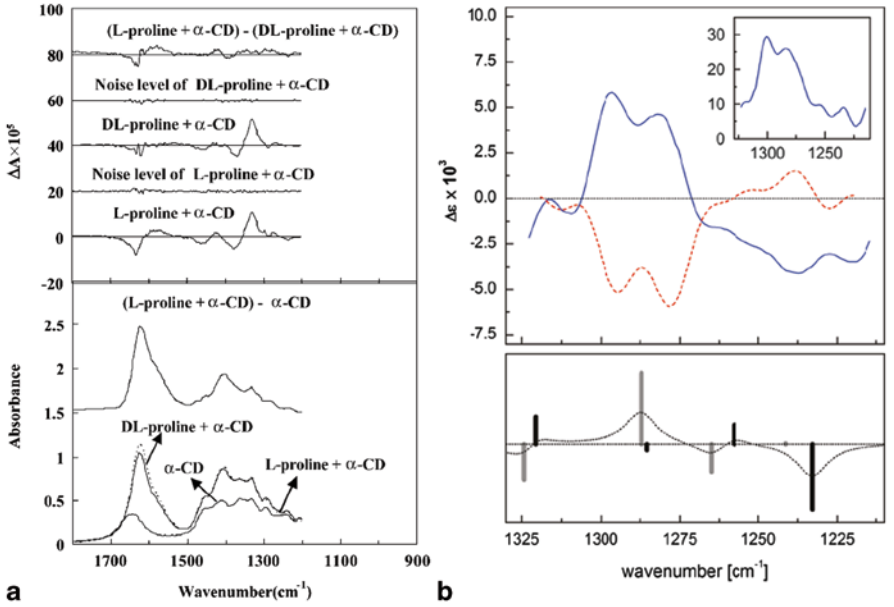


Fig. 5.8 (**a**) VCD and IR spectra of L-proline: α-CD film was prepared from a 200 μL solution of α-CD (0.050 M) in H_2O; DL-proline + α-CD film was prepared from a 200 μL solution of DL-proline + α-CD (molar ratio: 2:1, 0.050 M for α-CD) in H_2O; L-proline + α-CD film was prepared from a 200 μL solution of L-proline + α-CD (molar ratio: 2:1, 0.050 M for α-CD) in H_2O. (Reproduced from Ref. [38] with kind permission of Society of Applied Spectroscopy) (**b**) QCL-VCD spectra of aqueous 3 M solutions of L-proline (*solid blue*) and its enantiomer (*dashed red*) in a 100 μm CaF_2 cell. Spectra were background-corrected using a 1:1 (3 M) mixture of the two enantiomers. Inset: 3a in water corrected with a pure water background. b) Relative VCD intensities computed at the B3LYP/COSMO/6-31++G** level for two ring-puckering conformers of the L-proline zwitterion (1:1, grey and black vertical lines, arb. units). (Reproduced from Ref. [204] with kind permission of American Chemical Society)

studies [38, 149] and with the support of B3LYP/COSMO/6-31++G** calculations, the bands at 1295 and 1286 cm^{-1} of opposite sign were assigned to two methine bending vibrations of the two ring-puckering conformers (Fig. 5.8b) [204].

ROA spectra: Experiment and Calculations The first ROA and Raman spectra of L- and D-proline zwitterionic forms were recorded for light and heavy water solutions in a wide frequency range and interpreted by means of B3LYP/COSMO/6-31++G** calculations supported by HF, MP2, and other DFT functional calculations involving both a continuum and an explicit solvent (Fig. 5.9) [206]. The positions of the water molecules around proline were found by MD simulations. For each of the two proline conformations, 13 complexes with water (with from 3 to 8 H-bonded waters) were arbitrarily selected for simulation of the spectra. The 6-31G (water) and 6-31++G** (proline) basis sets were used for HF calculations of the intensity tensors. Two distinct but equally populated flexible

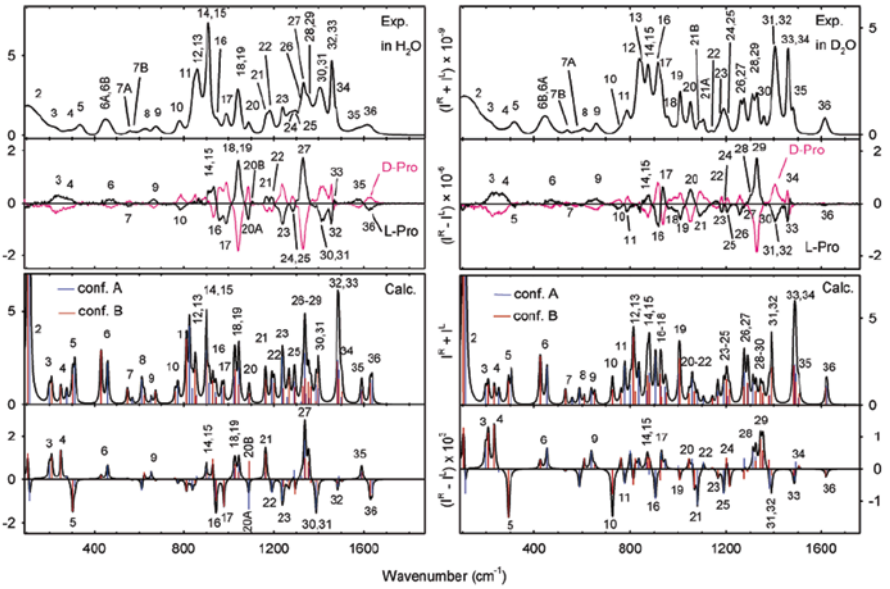


Fig. 5.9 Experimental (*top*) and calculated (*bottom*) Raman and ROA spectra of the L-proline zwitterion (*left*) and ND_2 deuterated L-proline zwitterion (*right*) in aqueous solutions. The calculations (L-enantiomers) were performed at the B3LYP/6-31++G**/COSMO level for both conformers surrounded by three to eight water molecules; band positions are marked by the blue and red lines for the **A** and **B** conformers, respectively, and the envelope was obtained as an average. (Reproduced from Ref. [206] with kind permission of American Chemical Society)

proline forms were found in solution (and glass). The band broadenings, reproduced based on dynamic averaging found from the 2D potential energy surfaces, revealed information about the extent of the proline ring motion and rotation of the carbonyl group [206]. Comparison of the computed and experimental bandwidths suggests that puckering is strongly correlated with carbonyl rotation. The study indicated that the environment modulates the properties of the hydrophobic part of the molecule indirectly by interacting with the ionic group. It was stressed that the experimental spectra could not be explained and interpreted without dynamic and solvent factors involved in the modelling.

As described for L-alanine, the role of internal motion in L-proline, ring puckering, rotation of the CO_2^- group, and anharmonic factors on the ROA spectrum were analysed by Daněček et al. [199]. Many DFT functionals, careful exploration of intensity tensors, degeneracy-corrected perturbation calculations, and large basis sets such as aug-cc-pVDZ and aug-cc-pVTZ (in conjunction with the B3PW91 functional) were exploited. Unlike alanine, for proline, the anharmonic averaged simulation provides a reasonable agreement with the experiment in the 1550–1205 cm^{-1} region, yielding more realistic band shapes, frequencies, and ROA intensities.

Recently, Qiu et al. published an experimental and computational study on the pH-dependency of L-proline ROA spectra in aqueous solution (Fig. 5.10) [207]. The

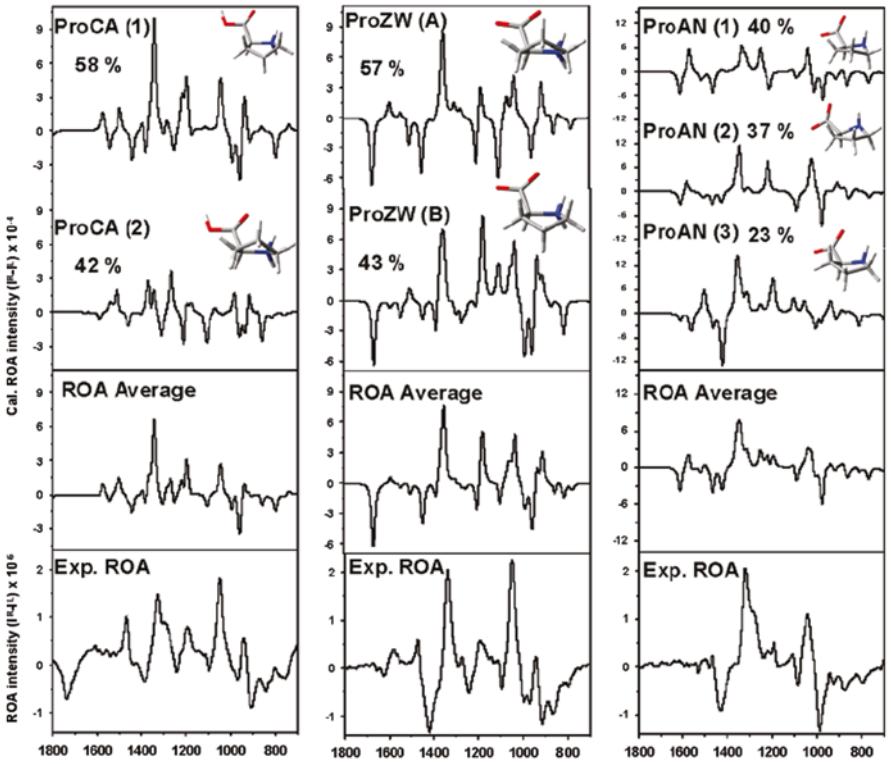


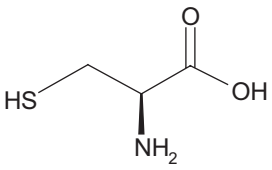
Fig. 5.10 ROA spectra of the two L-proline cation conformers (ProCA, *left*), two L-proline zwitterion conformers (ProZW, *middle*), and three L-proline anion conformers (ProAN, *right*), calculated at the B3LYP/PCM/aug-cc-PVDZ level, and the experimental spectra of L-Pro in 6 M HCl, H_2O, and 6 M NaOH. The Boltzmann population (in percent) and conformer shapes are indicated. Highly concentrated HCl and NaOH solutions are thought to completely suppress the zwitterion form. (Reproduced from Ref. [207] with kind permission of American Chemical Society)

B3LYP/PCM/aug-cc-pVDZ method without explicit water molecules was applied. The Raman shifts and ROA intensities were demonstrated to be directly correlated to the acidity and basicity of the solutions. A comparison of experimental and calculated spectra for the cation, zwitterion, and anionic forms showed differences in the bands corresponding to ring puckering and rotation of the carbonyl group. The noticeable changes in ROA intensities indicated that the acidic and basic environment acted not only by simple abstraction or addition of a proton but also changed the electron delocalisation throughout the molecule. Moreover, a more detailed analysis of the mode couplings suggested that the carbonyl group stretching modes might reflect both the electronic structure changes and the couplings between the three ring CH_2 group to the other groups contributing to the ROA intensities in the 1400–700 cm^{-1} range. Isolation of the two symmetric anisotropic invariants showed

the predominant contributions of the electric dipole–magnetic dipole invariant to the ROA signals in L-proline.

Finally, the ROA spectra of 4(*R*)-hydroxyproline in aqueous solutions at three different pH values were interpreted by B3LYP/IEF-PCM/aug-cc-pVDZ calculations for 10 anionic, 5 cationic and 15 zwitterionic conformations (Fig. 5.11) [208]. The ROA intensities were calculated at the HF/aug-ccpVDZ level, for which using the polarisable continuum model for calculations of optical tensors was found to be particularly problematic for the zwitterionic structures. As for proline, the endo forms of hydroxyproline appeared to be the most stable conformers. The ROA spectra calculated for the zwitterionic and cation forms stem predominantly from the two lowest conformations each, whereas for the isolated anion, the spectrum originated from one while in aqueous solution from four low-lying conformers. The largest ROA intensities, negative or positive, were assigned to stretching vibrations of the five-membered ring of 4-trans-(R)-hydroxyproline. No simple correlation between the ring conformation and the ROA intensity of those vibrations was found.

5.5.2.3 L-cysteine



Scheme 6 (S)-2-amino-3-mercaptopropionic acid

VCD and ROA: Experiment and Calculations The VCD C*H stretching and bending vibration bands of L-cysteine were registered and compared with those of more than 20 other amino acids in the 1980s [147–150]. Since then, the vibrational optical activity of L-cysteine has been studied only sporadically. In 2006, Pecul computationally studied the conformational dependence on ROA spectra for 23 stable neutral structures [209], showing the ROA intensities of cysteine to significantly change their sign and magnitude as the conformation changed. She found that strongly negative ROA ν(S-H) and ν(C-S) stretching vibration modes, when accompanied by a negative value for the δ(S-H) deformation mode, indicated that the SCCN dihedral angle was close to a straight angle, whereas when accompanied by positive ROA intensity for δ(S-H), the angle value was close to 60°. A year later, when discussing the measured Ar-matrix isolation IR spectra of (neutral) L-cysteine, Dobrowolski et al. demonstrated that calculated VCD spectra might be much more distinctive for the L-cysteine conformers than the IR spectra [161]. Also, in 2007, Sadlej et al. studied the VCD spectra of more than forty 1:1 complexes of neutral L-cysteine and water at the B3LYP/aug-cc-pVDZ level [210]. They showed

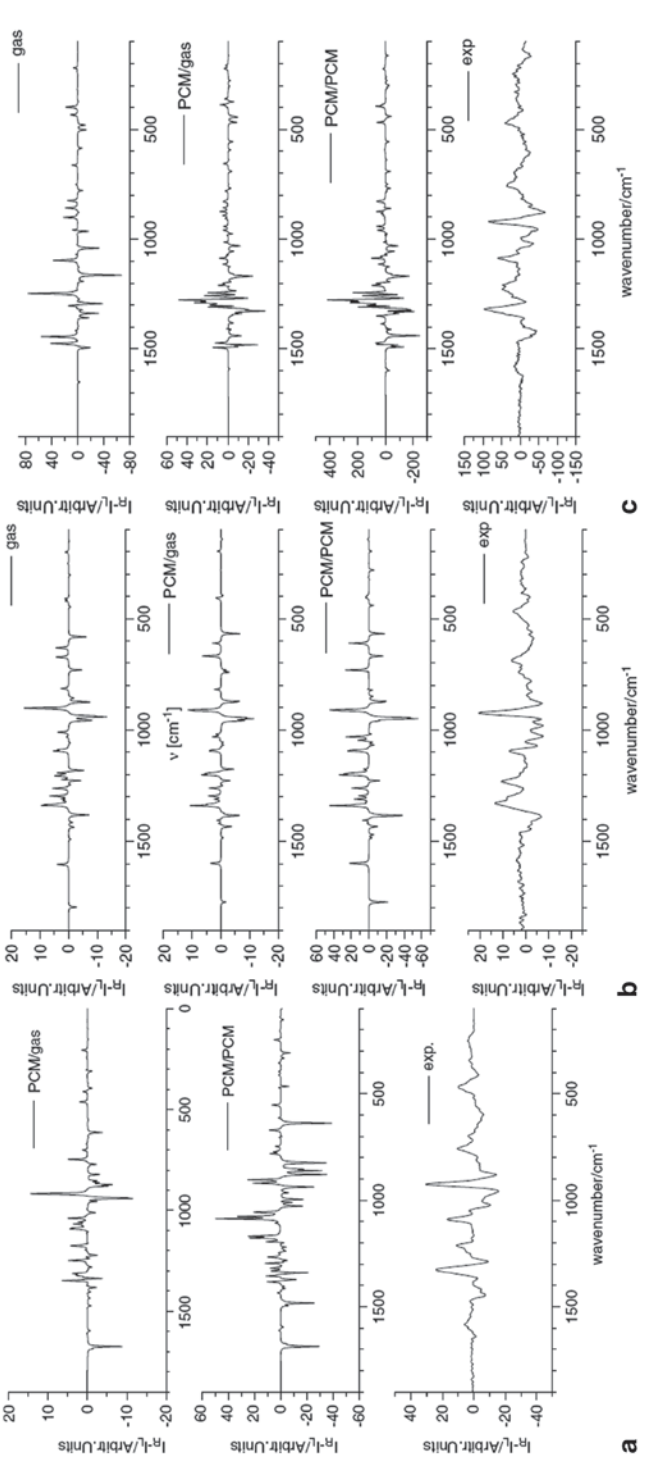


Fig. 5.11 The calculated Boltzmann-averaged ROA of the zwitterion (**a**), cationic (**b**), and anionic (**c**) form of 4(*R*)-hydroxyproline (in arbitrary units). The experimental ROA spectrum was recorded for ca. 100 mg/ml of amino acid (**a**) at pH 6, acquisition time: 5679s, power at sample 598 mW; (**b**) at pH 1, acquisition time: 3523s, power at sample 524 mW; (**c**) at pH 10, acquisition time: 3993s, power at sample 500 mW. (Reproduced from Ref. [208] with kind permission of John Wiley and Sons)

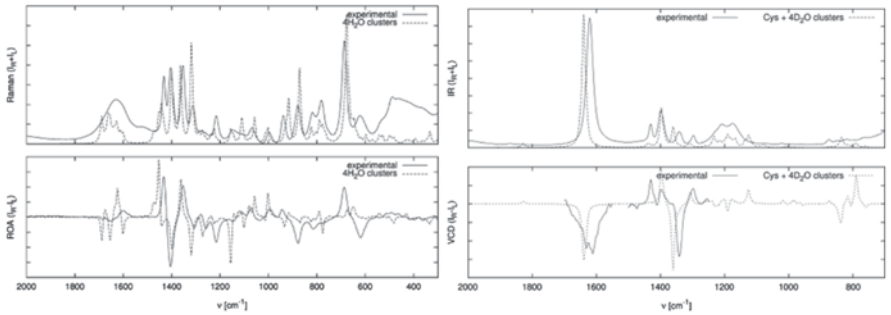


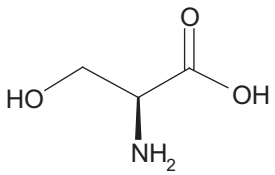
Fig. 5.12 (**a**) Comparison of the Raman (*upper* graph) and ROA (*lower* graph) spectra obtained from optimised L-cysteine-water clusters by experiment. (**b**) Comparison of the IR (*upper* graph) and VCD (*lower* graph) spectra obtained from optimised L-cysteine-water clusters by experiment. (Reproduced from Ref. [140] with kind permission of American Chemical Society)

that the ν(OH) and ν(NH_2) and β(OH_2) and β(NH_2) ranges of the VCD spectra were especially useful for discriminating between different chiral forms of intermolecular hydrogen-bonding complexes. In particular, they characterised the VCD chirality transfer phenomenon [67, 211, 212] occurring from chiral L-cysteine to an achiral water molecule after complex formation in which the water molecule acquired significant rotational strengths whose signs changed in line with the geometry of the complex. They suggested ν^a(NH_2), β(NH_2), and ν(C=O) to be relatively insensitive to complex structural changes.

Very recently, Kamiński et al. measured and thoroughly interpreted the VCD, ROA, IR, and Raman spectra of L-cysteine in aqueous solution (Fig. 5.12) [140]. The interpretation was done by means of the DFT/aug-cc-pVDZ method combined with either the IEF-PCM model or a microsolvation model (four water molecules) obtained from molecular dynamics modelling. The authors confirmed previous findings, indicating that although use of the sole polarised continuum model reproduced some of the features of the L-cysteine zwitterion spectra in water, the best description was rendered by the solute-solvent clusters embedded in bulk water represented by a PCM model. They found that MD yielded unsatisfactory band positions and intensities, and, even if in some cases it more correctly reproduced band shapes, the standard force field parametrised for peptides cannot be recommended for studies of isolated amino acids. Interestingly, the authors found the conformer population from fitting the experimental ROA spectra by linear combination of the computational spectra. Based on this procedure, three zwitterion conformations were found to be present at 35, 33, and 24 %. The spectral range in which the experimental VCD spectrum of L-cysteine in water could be registered was narrower than in the ROA, IR, or Raman spectra (Fig. 5.12b). Based on a complex experimental and theoretical study, a cautious and critical assignment of the vibrational bands of L-cysteine in aqueous solution was provided for all four types of vibrational spectra.

Let us mention that the VCD, IR, and UV-Vis spectra of N-acetyl-L-cysteine, a drug used in detoxifying free radicals, were also measured in aqueous solution (D_2O) at several pH levels [137]. The spectra were interpreted by means of B3LYP/6-311++G** calculations and molecular dynamics simulations. As found previously for unsubstituted L-cysteine [137], the energetic order of the protonated, neutral, mono- and doubly-deprotonated N-acetyl-L-cysteine conformers was noticeably changed after the inclusion of the polarisable continuum model. To properly account for the discrepancies noted between the experimental and simulated VCD spectra, the PCM and the explicit solvent model (preceded by the MD simulations of N-acetyl-L-cysteine water clusters) were used. The IR and VCD intensities computed for N-acetyl-L-cysteine clusters with three or four water molecules at the B3LYP/6-311++G** level with and without the PCM indicated that combination of explicit and implicit solvation models provided better agreement between theory and experiment, allowing for finding the conformational distributions of N-acetyl-L-cysteine in water and hydrogen bonding interactions between the solute and water molecules.

5.5.2.4 L-serine



Scheme 7 (S)-2-amino-3-hydroxypropanoic acid

VCD and ROA: Experiment and Calculations The CH region of the VCD spectra of L-serine was registered in saturated water solutions and for solid samples, as mulls in halocarbon oil, at the end of the 1970s [169]. Subsequently, the VCD C*H stretching and bending vibration regions of L-serine and other amino acids were studied in the 1980s [147–150]. Since then, the VCD spectrum of serine in water has been the subject of only one very recent systematic study by Zhu et al. [213]. A careful collection of VCD and IR spectra in aqueous solutions under three pH conditions of 1.0, 5.7, and 13.0 in light and heavy water enabled the authors to provide a detailed description of the conformational distributions and VCD and IR features of the dominant protonated, zwitterionic, and deprotonated species of serine (Fig. 5.13). B3LYP/6-311++G** calculations and molecular dynamics simulations were used to support the interpretation. Presence of six explicit water molecules around the L-serine species, found from radial distribution function analyses and molecular dynamics snapshots, was found to noticeably improve the agreement between the computational and experiment spectra of each of the serine species.

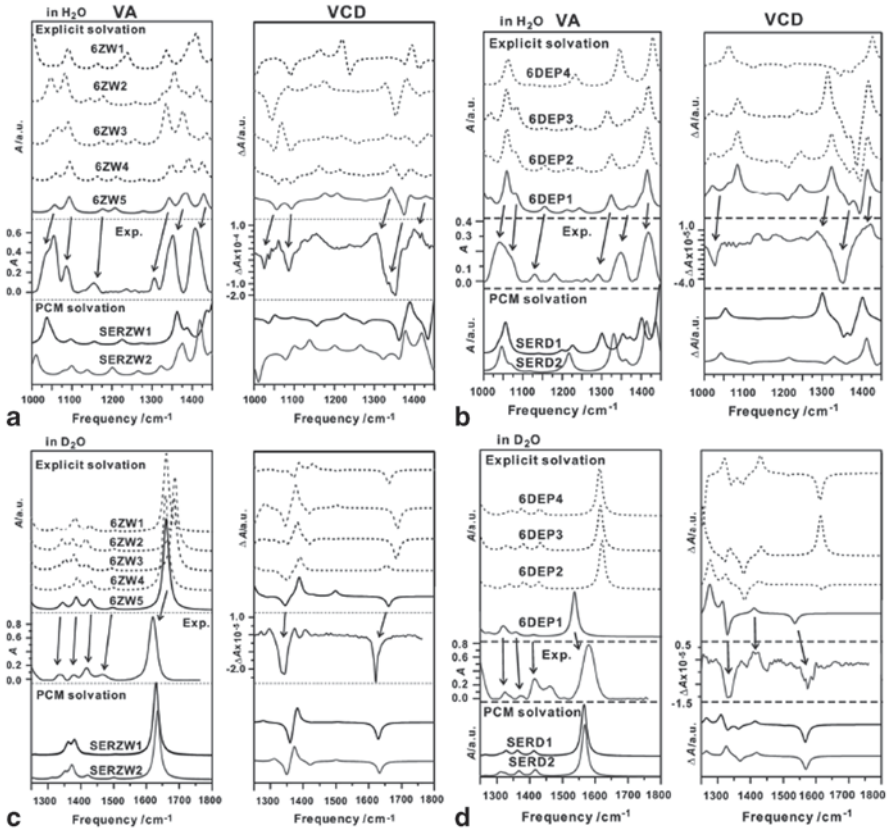


Fig. 5.13 *Left*. Comparisons of the simulated VA and VCD spectra of the five most stable zwitterionic serine–$(water)_6$ conformers with the related experimental data at neutral pH in water (a) and in D_2O (c). The corresponding PCM simulated VA and VCD spectra of SERZW1 and SERZW^2 in water and in D_2O are depicted at the bottom. *Right*. Comparisons of the simulated VA and VCD spectra of the four most stable deprotonated serine–$(water)_6$ conformers with the related experimental data at pH 13 in water (b) and D_2O (d). The corresponding PCM simulated VA and VCD spectra of the deuterated SERD1 and SERD2 in water and D_2O are depicted at the bottom. (Reproduced from Ref. [213] with kind permission of John Wiley and Sons)

The ROA spectra of neutral L-serine conformers in water were calculated by Pecul in 2006 [209]. As for cysteine, the main aim of the study was to relate the changeability of the ROA spectrum with structural variations occurring as the serine conformation changed. However, unlike for cysteine for which interactions of side SH groups are relatively weak, the formation of hydrogen bonds by the side OH group was found to be the deciding factor determining the ROA intensities of the side chain vibrations [209].

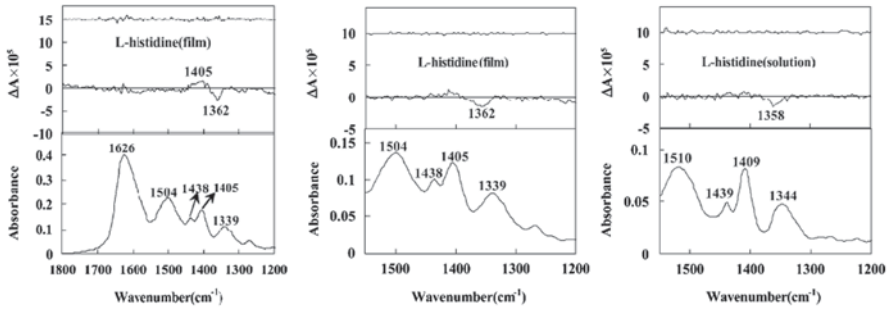
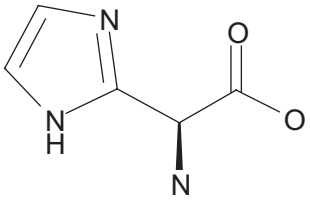


Fig. 5.14 A film of L-histidine was obtained from a 200 μL solution of L-histidine-α-CD (0.012 M) in H_2O. The solution spectra of L-histidine were obtained at 0.26 M using a 15 μm path length demountable cell. Solution VCD spectra were obtained with 6 h of data collection time. (Reproduced from Ref. [38] with kind permission of Society of Applied Spectroscopy)

5.5.2.5 L-histidine



Scheme 8 (S)-amino-(1H-imidazol-2-yl)-acetic acid

VCD and ROA: Experiment and Calculations Beyond the early registrations of C*H band vibrations in VCD spectra by the Nafie group [147–150], the VCD spectra of L-histidine in water have been collected in water solution and in the form of aqueous films with α-cyclodextrin (Fig. 5.14) [38].

The thorough computational investigations of Jalkanen et al. [154] on the vibrational spectroscopy of three amino acids in aqueous solution also devoted some attention to L-histidine. The authors demonstrated that, as for alanine, the supramolecular treatment of the first solvation sphere by at least four water molecules was necessary to stabilise L-histidine zwitterions. The mono- and di- protonated histidine zwitterions were considered to reproduce the VCD and ROA spectra.

A combined theoretical and experimental VCD, ROA as well as IR and Raman study of L-histidine in aqueous solution was done by Deplazes et al. in 2008 [214]. Based on B3LYP/PCM/6-31G* calculations performed with the use of 13 water molecules, the authors verified previous findings regarding the kind and conformation of species present in the solutions. The verification was based on absorption IR and Raman spectra of L-histidine cations, zwitterions, and anions measured at various pH levels in aqueous solution, and then applied to reproduce the VCD and ROA intensities (Fig. 5.15).

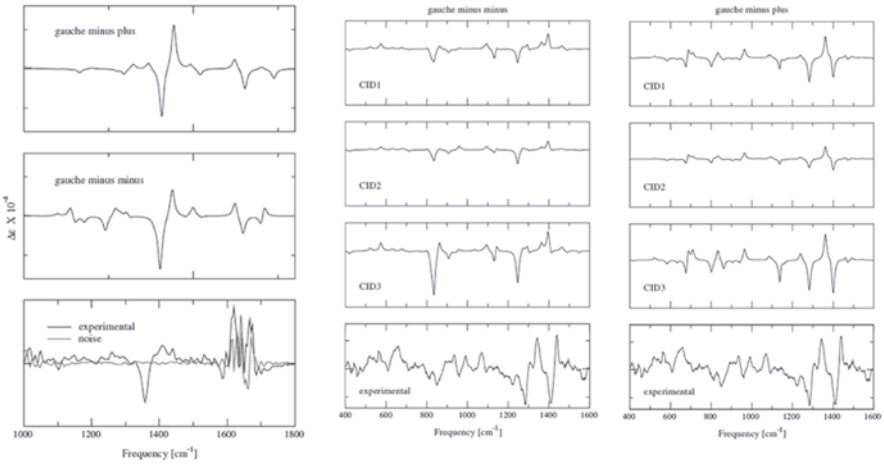
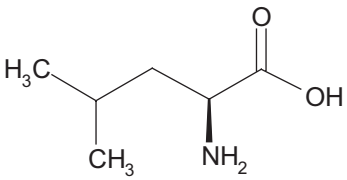


Fig. 5.15 Comparison of B3LYP/PCM/6-31G* calculated VCD (*left*) and ROA (*middle* and *right*) of the L-histidine GMM (*middle*) and GMP (*right*) zwitterions solvated with 13 explicit water molecules with the experimental spectrum (*bottom*) of an L-histidine water solution. (Reproduced from Ref. [214] with kind permission of Springer Science + Business Media)

5.5.2.6 L-leucine



Scheme 9 (S)-2-amino -4-methylpentanoic acid

VCD and ROA: Experiment and Calculations The VCD C*H stretching and bending vibration bands of L-leucine in water were observed in early studies by the Nafie group [147, 149]. Since then, no report on L-leucine VCD spectra was made until the very recent study performed by the Xu group [215–217]. The IR spectra of L-leucine in H_2O and D_2O were recorded under three representative pH conditions (1.0, 6.4, and 13); however, because of the poor solubility of L-leucine in strongly acidic or basic solutions, the VCD spectra were taken only at near-neutral pH/pD. As for L-serine and N-acetyl-L-cysteine in aqueous solutions [137, 213], the conformational distribution of the dominant protonated, deprotonated, and zwitterionic leucine species in water was studied in detail using quantum chemical calculations (B3LYP/6-311++G(d, p)). Additionally, the molecular dynamics simulation of a leucine molecule solvated with 620 water molecules allowed for the rational selection of zwitterionic-$(water)_n$ clusters for further DFT studies. The authors found three protonated, three deprotonated, and three zwitterionic forms.

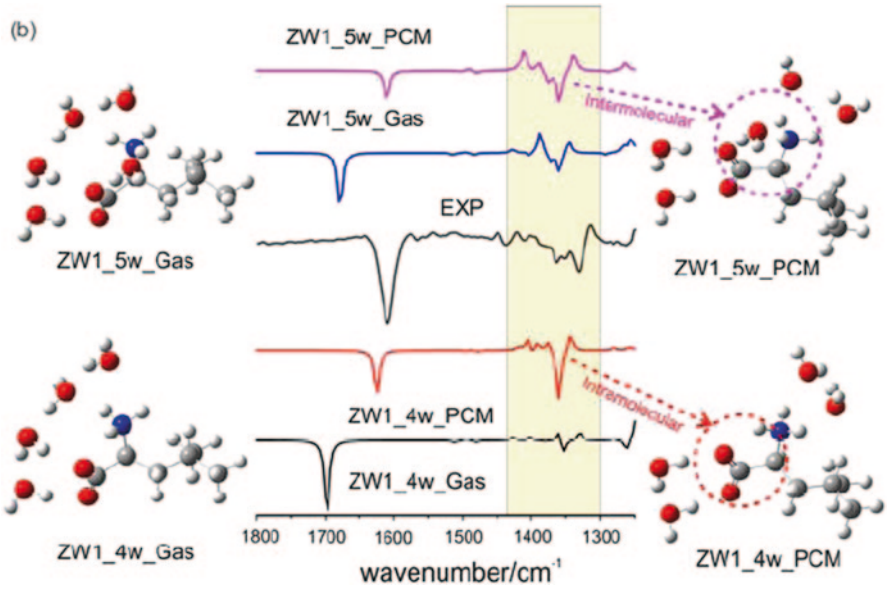


Fig. 5.16 VCD spectra of L-leucine in neutral aqueous (D_2O) solution: measured (**a**) and (**b**) predicted for different conformations and their superposition (in black). B3LYP/PCM/6-311++G(d, p) calculations were applied (Reproduced from Ref. [216] with kind permission of American Institute of Physics)

It appeared that, in each case, one form accounted for at least 70% abundance in the mixture when the implicit IEF-PCM solvation model was used. Inclusion of the explicit water molecules to improve the agreement between theory and experiment is under study. This is important because, indeed, it was shown that the stability of the zwitterionic over the neutral form of leucine, modelled at the CPCM/B3LYP/6-311++G(d, p) level, increased from 1.6 kcal/mol to 4.8 kcal/mol for zero and two water molecules, respectively [217]. The IR spectra at the three pH levels were noticeably different because of the existence of different dominant species at each pH. The differences were easily demonstrable by the position of the C=O stretching vibration band, the location of which varied by over 150 cm^{-1} from 1723 cm^{-1} (neutral), to 1611 cm^{-1} (zwitterionic), to 1563 cm^{-1} (deprotonated). Interestingly, below 1800 cm^{-1}, the absorption IR spectra of leucine in water can be satisfactory reproduced by the sole PCM solvation model. This is, however, not the case for the leucine VCD spectrum for which the PCM-solvated leucine-$(water)_n$ cluster (n=4,5) was necessary for a fair reproduction of the experimental spectra (Fig. 5.16). The complex experimental spectral pattern in the 1500–1300 cm^{-1} region was best captured by the leucine-$(water)_5$/PCM cluster in which one water molecule was locked between the COO^- and the NH_3^+ groups. In conclusion, the authors stated that to faithfully account for the solvent effects of leucine in water and to reproduce the VCD spectra, the inclusion of the implicit solvent model combined with explicit water solvated clusters is critical.

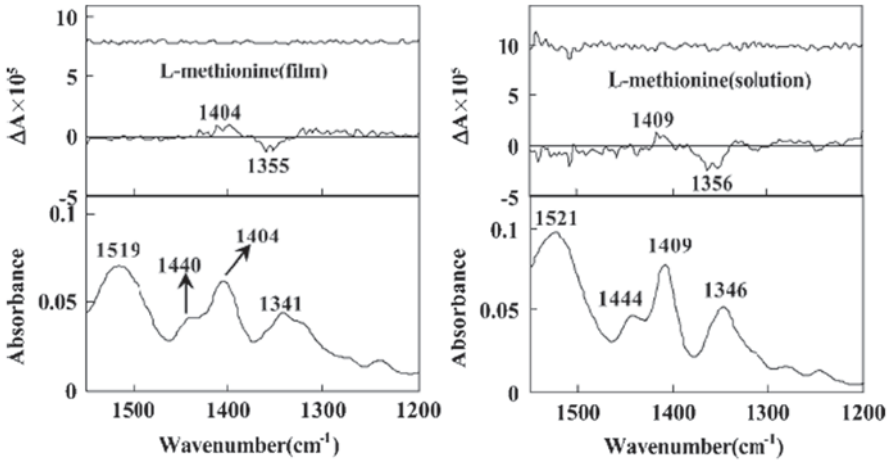
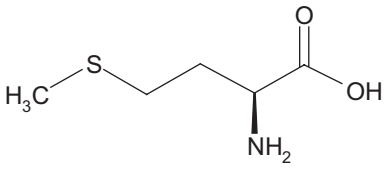


Fig. 5.17 A film of L-methionine was obtained from a 400 μL solution of L-methionine-**α-CD** (0.014 M) in H_2O. Solution spectra of L- methionine were obtained at 0.20 M using a 15 μm path length demountable cell. Solution VCD spectra were obtained with 6 h of data collection time (Reproduced from Ref. [38] with kind permission of Society of Applied Spectroscopy)

5.5.2.7 L-methionine



Scheme 10 (S)-2-amino(methylthio) butanoic acid

VCD and ROA: Experiment and Calculations So far, for L-methionine, the only VCD measurements have been done by the Nafie group in the 1980s [147–150], and by Zhang and Polavarapu in films with α-cyclodextrin (Fig. 5.17, Ref. [38]).

Very recently, Rode et al. found 23 methionine zwitterions to be stable in water, simulated by the PCM model at the B3LYP and MP2 levels combined with the aug-cc-pVDZ basis set [218]. The conformer abundance appeared to be very sensitive to the calculation level. However, at the MP2 level, structures with two intramolecular hydrogen bonds (IHB), i.e., NH···S-, = and NH···O=, dominate (three structures populated more than 5% were predicted). Further, microsolvation of methionine zwitterions by five water molecules was taken into account for predicting the VCD spectra. Five water molecules in the first solvation sphere guarantee that for structures with two internal H-bonds (IHBs), all electron-donor and electron-acceptor sites of methionine are saturated. For ten methionine zwitterions (with and without two IHB), different clusters with $(H_2O)_5$ were considered at the B3LYP-IEF-PCM

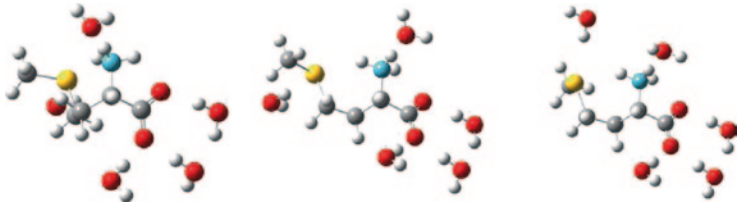


Fig. 5.18 The Methionine zwitterion clusters with $(H_2O)_5$ populated according to the ΔG values obtained at the B3LYP-IEF-PCM/aug-cc-pVDZ level. [218]

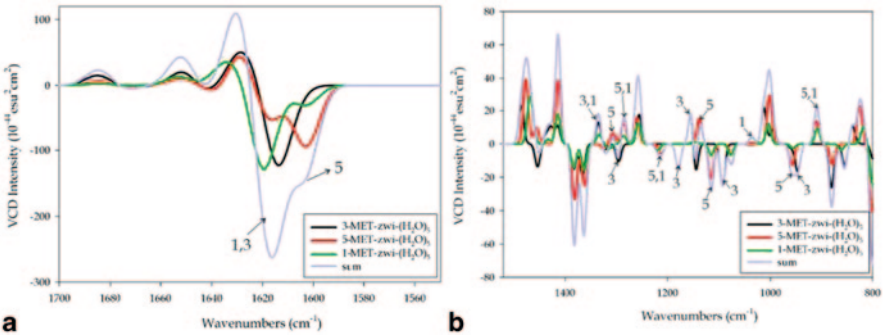


Fig. 5.19 The simulated spectra of methionine in water (*light-violet*) calculated at the B3LYP-IEF-PCM/aug-cc-pVDZ level. [218]

level. It was shown that the population of these complexes strongly depends on the inclusion of the thermochemical term. The formation of strong intermolecular hydrogen bond complexes with water increases the entropy term, and the structures favoured in the ΔE scale, without the thermochemical correction, are definitely different from the ΔG favoured ones. Based on the $\Delta G_{B3LYP\text{-}IEF\text{-}PCM}$ values, three MET-zwi-$(H_2O)_5$ clusters are populated more than 5% (Fig. 5.18). All these complexes exhibit two IHB, i.e. the NH···O= and NH···S however, the structures differ from each other by at least one methionine dihedral.

The simulated VCD spectra enable the recognition of these three conformers in their mixture (Fig. 5.19). Interestingly, the VCD intensities of the least populated 1-MET-zwi-$(H_2O)_5$ complex are very strong, enabling the assignment of its bands in the weighted VCD spectrum. The chirality transfer VCD bands, i.e., scissoring bending water modes δ(H_2O), occur in the 1700-1550 cm^{-1} range. The most intense band in the entire VCD spectrum corresponds to the δ(H_2O) mode coupled with the ν(C=O) stretching mode of the carboxylic group.

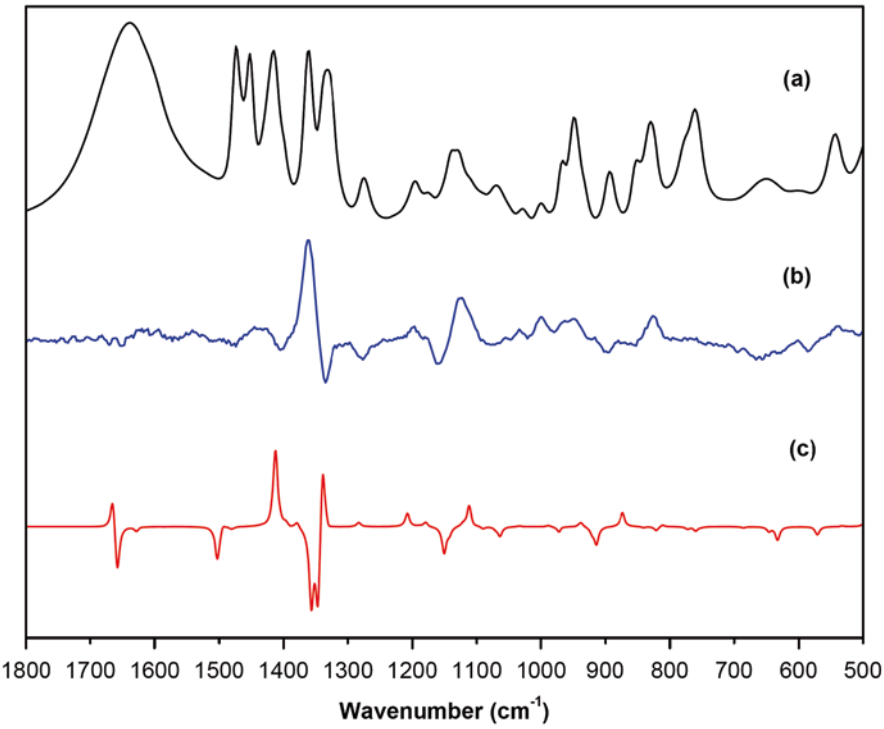
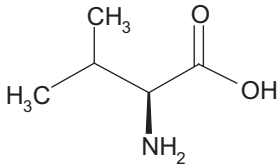


Fig. 5.20 The comparison between Raman and Raman Optical Activity experimental and calculated (B3LYP/aug-cc-pVTZ, PCM) spectra of L-valine (**a**- Raman experimental spectrum, **b**- Raman Optical Activity experimental spectrum, **c**- Raman Optical Activity theoretical spectrum). [219]. Reproduced with kind permission of the authors.)

5.5.2.8 L-valine



Scheme 11 (S)-2-amino -3-methylbutanoic acid

VCD and ROA: Experiment and Calculations The VCD spectra of L-valine have been measured only by the Nafie group [147–150], whereas the ROA spectra of L-valine in aqueous solution at pH 7–9 were recently measured in the 2000–500 cm^{-1} range with a spectral resolution of 7 cm^{-1} (Fig. 5.20) [219]. The ROA spectra were interpreted by means of B3LYP/PCM/6-311++G** and B3LYP/PCM/aug-cc-pVDZ calculations for the L-valine zwitterion in aqueous solution.

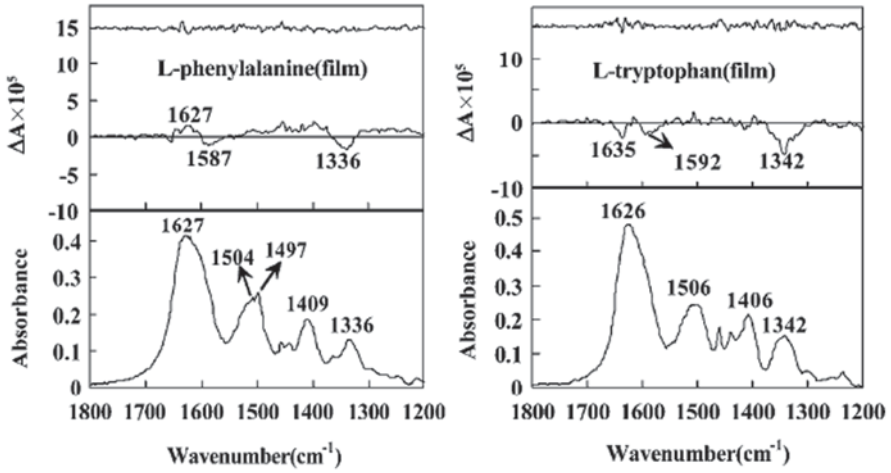
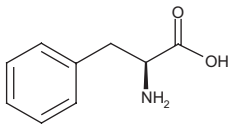
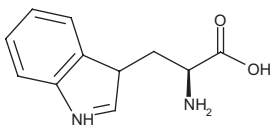


Fig. 5.21 A 200 μL solution of L-phenylalanine-α-CD (molar ratio: 1:1, 0.024 M), and a 420 μL solution of L-tryptophan-α-CD (molar ratio: 1:1, 0.019 M) in H_2O were used, respectively, to obtain the VCD spectra of these amino acids in the film state. (Reproduced from Ref. [38] with kind permission of Society of Applied Spectroscopy)

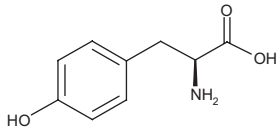
5.5.2.9 L-phenylalanine, L-tryptophan, L-tyrosine



(S)-2-amino-3-phenyl propanoic acid



(2S)-2-amino-3-(1H-indol-3-yl)propanoic acid



(S)-2-Amino-3-(4-hydroxyphenyl) propanoic acid

Scheme 12 L-phenylalanine, L-tryptophan, L-tyrosine

VCD and ROA: Experiment and Calculations As for many other amino acids, for L-phenylalanine, the only VCD measurements were done for the C*H band stretching and bending vibration bands by the Nafie group in the 1980s [147–150], and in the form of a film with α-cyclodextrin by Zhang and Polavarapu [38]. For L-tryptophan, the only VCD measurements were done in a film with α-cyclodextrin (Fig. 5.21) [38].

L-tryptophan was also considered in the complex computational investigations of the VCD and ROA spectra of amino acids and peptides by Jalkanen et al. [154]. It was found that with the use of the cc-pVDZ basis set, PBE and PW91 functionals, and the Onsager model of the aqueous environment, the zwitterionic structure of L-tryptophan was not stable, whereas it became stable when diffused functions were added in the form of the aug-cc-pVDZ basis set. In this paper, in the case of the L-tryptophan molecule, the authors were mainly focused on methodological problems connected with the best choice of basis set or functional to reproduce the zwitterion's stability.

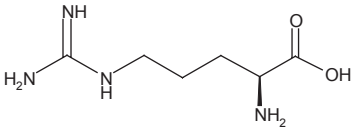
In 2008, Jacob et al. discussed a ROA signature yielded by tryptophan in the peptide backbone based on the BP86/TZVP calculated spectra of six different conformers of N-acetyl-(S)-tryptophan-N'-methylamide [220]. They confirmed the earlier experimental findings [221] that the ROA intensity of the tryptophan band at ca. 1540 cm^{-1} reflected the conformation of the tryptophan side chain. In 2009, Luber and Reiher described a method for the quantum-chemical determination of modes with maximum Raman and ROA intensity (intensity-tracking calculations) applied for several chiral molecules including L-tryptophan [222]. The authors suggested a method to reveal especially high potential for large molecules, for which ordinary calculations are very time-consuming. The procedure, applied to L-tryptophan, selected the band at ca. 1540 cm^{-1} as the most appropriate for tracking the molecule.

Neither VCD nor ROA spectra have been registered for L-tyrosine in aqueous solution because of its extremely low solubility even in the film state [38]. The only computational study done for this molecule was that performed by Tanaka et al. in 2006 mimicking the spectra of aromatic amino acid residues in a protein [165].

5.6 Summary and Perspectives

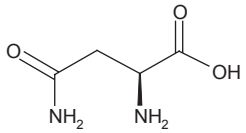
5.6.1 Not-yet-measured Amino Acids

L-arginine



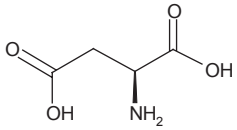
(S)-2-amino-5-guanidinopentanoic acid

L-asparagine



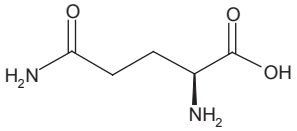
(S)-2-amino-3-carbamoylpropanoic acid

L-aspartic acid



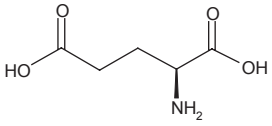
(S)-2-aminobutanedioic acid

L-glutamine



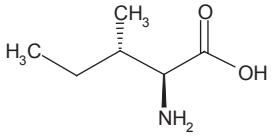
(S)-2-amino-4-carbamoylbutanoic acid

L-glutamic acid



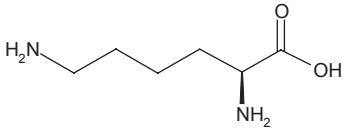
(S)-2-aminopentanedioic acid

L-isoleucine



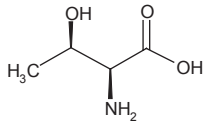
(2S,3S)-2-amino-3-methylpentanoic acid.

L-lysine



2(S),6-diaminohexanoic acid

L-threonine



2S,3R-hydroxybutanoic acid

Of the other natural amino acids, the VCD spectra of asparagine, arginine hydrochloride, glutamine, isoleucine, leucine, lysine hydrochloride, threonine, allothreonine, and valine were measured only in the 1980s by the Nafie group [147–150], whereas it seems that those of aspartic and glutamic acids have not yet been registered.

5.6.2 Challenges for Computational Methods

In this Chapter, we briefly reviewed the application of theoretical studies to understand the relationship between the conformation of chiral molecules and experimental VCD and ROA spectra. Thanks to the introduction of the analytic schemes for the calculation of geometric derivatives of second-order molecular tensors, routine calculations of spectra are possible, although the computational cost remains high, especially for ROA spectra. However, several challenges still exist in computational chiroptical spectroscopy. We would like to mention some major future directions:

1. state-of the-art choice of the method of calculation;
2. the problem of conformational flexibility of molecules;
3. the effects of solvent on the frequency and the intensity of chiroptical properties.

Recently, the coupled clusters (CC) approach has been applied to chiral molecules containing up to 30 to 40 atoms. However, the level of dynamic electron correlation and mechanical/electric anharmonicity needed to reproduce gas-phase measurements is still unclear. Also, the need to calculate geometric derivatives of the tensors **α**, **G'** and **A** with respect to the nuclear coordinates remains active. Moreover, the problem of basis set completeness at the CC and DFT levels of theory also belongs to the first point, especially because the requirements of the basis set for ROA intensity calculations are more demanding than for VCD calculations.

The second point is related to biomolecules (amino acids) which can exist in several conformations. The number of conformations increases rapidly with molecular size. Thus, the difficulty in the location of relevant stable structures increases rapidly as well. Different conformations can have very different VCD and ROA spectra; therefore, for these conformationally flexible molecules, the use of the Boltzmann averaging procedure in the calculation of the spectra is necessary. This means that the relative energy of different conformations should be obtained with high accuracy. The question of whether to use the electronic energy differences (ΔE) or the Gibbs energy differences (ΔG) remains open. Even though the zero-point vibrational energy (ZPE) is rarely considered, it may be necessary to get the correct sign of several chiroptical properties [223].

The measurement of biomolecules in a water environment is one of the main benefits of ROA over the VCD technique. However, solute-solvent interactions, such as hydrogen bonds, cannot be properly described in the frame of the sole PCM models. Therefore, explicit solvation models are needed to account for the influence of the environment on chiroptic spectra. Accurate sampling of the solute-solvent

configurational landscape and its dynamics is currently a challenging aspect of computational work.

5.6.3 Promising Experimental Techniques

Measuring vibrational chiroptical spectra meets two serious problems of an objective nature: (1) the sensitivity of the VCD and ROA spectra and (2) the very low solubility of several amino acids in water. Continuous improvement of spectrometers facilitates and ameliorates the quality of VCD and ROA measurements. However, measurements of the VCD spectra with a tunable external-cavity quantum cascade laser (QCL), as recently done by Lüdeke et al. for proline [204], are especially promising. QCLs provide at least 100 times more power in the MIR than ordinary thermal IR light sources and permit the use of IR detectors that do not require N_2 cooling. A shortcoming of the method is in the very limited wavenumber range of ca. 100 to 200 cm^{-1} in the 800 and 2500 cm^{-1} MIR interval. However, even now, this spectral range would permit very valuable measurements of amino acid VCD spectra.

Another promising innovation may be the introduction of a transient vibrational chiral spectrometer, capable of detecting photo-induced VCD changes with picosecond time resolution, as recently described by Bonmarin and Helbing [224–228]. With this newly developed spectrometer, it was possible to detect VCD changes in the CH-stretch region following visible excitation of a cobalt spartein complex ($Co(sp)Cl_2$) with picosecond time resolution [224]. Also, it was possible to significantly increase the signal to noise ratio of the static VCD and VORD spectra and ameliorate the quality of the transient chiral signals. It is expected that this technique will make it possible to identify chiral intermediates in the course of a chemical reaction or to probe with unprecedented detail the conformational dynamics of amino acids and local secondary structure formation in peptides and proteins [224].

Yet another technique which essentially extends the possibilities of VCD spectroscopy is application of the two-dimensional VCD (2D VCD) method, known for more than a decade [184] but still rarely applied [176, 184–186]. As demonstrated by Ma et al. for alanine in a water solution, 2D VCD enables significant resolution enhancement of overlapping bands and assignments of combination bands, especially in the NIR VCD region [176]. A variety of 2D spectra can be registered in synchronous and asynchronous modes, providing the possibility of combination of homo- and hetero-region spectra of both IR-IR, IR-VCD, and VCD-VCD techniques [176]. This seems to provide an extremely broad potential for applications of 2D VCD spectroscopy in physico-chemical studies of amino acids and other important biomolecules.

The α-cyclodextrin film technique introduced to VCD spectroscopy by Shanmugam and Polavarapu [151], and then applied by Zhang and Polavarapu [38, 153] and combined with modern VCD spectrometers [55], is worth further exploration, since

it would provide new valuable results for poorly soluble amino acids. Some other innovations can be found in Ref. [56, 229, 230].

Water is a natural solvent for Raman measurements at different pH levels. Moreover, the array of Raman spectroscopy methods is much larger than those available for absorption IR spectroscopy. Therefore, we expect much more intensive studies on amino acids in water media by ROA than by VCD techniques. In the late 1990s, Nafie proposed the single electronic-state limit of the theory of Resonance ROA (RROA) effect and explained why the sign of all RROA bands is the same [231]. Subsequently, the RROA was measured [232–234] and the theory was extended and improved [235–237]. Quite recently, RROA involving more than one resonant electronic state has been reported [238]. Also very recently, induced resonance Raman optical activity (IRROA) has been detected in the presence of a europium complex, achieving a 10^4-fold increase in sensitivity in comparison to non-resonant ROA [235]. Registration of RROA of amino acids may be important for aromatic phenylalanine and tryptophan, and, especially, for tyrosine which, because of its extremely poor solubility in water, have not yet been studied by VCD and ROA methods.

Using a UV frequency laser for ROA measurements was considered, as changing the beam frequency from 488 to 244 nm would increase the ROA signals by a factor of 32 [55]. Also, the UV incident beam makes the near resonance Raman conditions additionally probable by increasing the intensity. Recently, Kapitán, Barron, and Hecht experimentally devised an ROA spectrometer with UV excitation wavelength of 244 nm [239]. This technique is promising for a significant extension of the capabilities of ROA spectroscopy. This would be especially interesting for amino acids and peptides, since indeed high-quality UV ROA spectra of zwitterionic cyclo(L-Ala-L-Ala) and cyclo(D-Ala-D-Ala) and other alanine peptides have been demonstrated [239]. On the other hand, excitation in NIR decreases the fluorescence of the sample. In 2010, it was reported that a near IR ROA spectrometer was constructed and applied for measurements of S-(-)-α-pinene and L-alanyl-L-alanine in water [240]. Also in 2010, NIR ROA measurements were performed on proteins using a 785 nm NIR diode laser [241]. Despite the fact that, with the NIR beam frequency, ROA signals are decreased by significant factor, such spectrometers seem to present high potential for measuring biomolecules *in situ* and, in particular, for measurements of free amino acids in their native surrounding.

In 2008, Abdali and Blanch surveyed the theoretical and experimental work undertaken to develop surface-enhanced Raman optical activity (SEROA), which is complementary to both SERS and ROA [242]. The application of SEROA spectra to studies of amino acids was done in 2011 for L-phenylalanine and L- and D-cysteine adsorbed on silver [243]. However, for phenylalanine, the presence of SEROA signals was questionable, while for L- and D-cysteine on a silver electrode mirror image surface, the ROA spectra were registered in a broad wavenumber region, confirming the observation of the SEROA effect. Theoretical models of the SEROA are continuously being developed [244–246]. Moreover, observations

of surface-enhanced resonance Raman optical activity (SERROA) have also been reported [234, 240, 247–249]. It is very probable that the new SEROA techniques will serve not only as methods yielding spectral fingerprints for different types of biological molecules and as monitors of chiral surface-analyte interactions, but will find valuable applications in physico-chemical studies on amino acid-metal surface interactions.

Finally, there has been significant progress in ROA measurements enabling simultaneous acquisition of all four forms of circular polarisation Raman optical activity. This progress is especially important for dual circular polarisation in-phase ROA, the most efficient form of ROA that will be important, particularly in biological applications [250].

Acknowledgments This work was partially supported by ICRI (Warsaw) within the framework of statutory activity. Partial support was also provided by Grant No. N N204 443 140 of National Science Centre in Poland. The computational G19-4 grant from the Interdisciplinary Center of Mathematical and Computer Modeling (ICM) of the University of Warsaw is also acknowledged.

Appendix

In Table A1, the amino acids together with their precursors are presented. Apart from their well-known role as building blocks of proteins, the α-amino acids can form some lipids, act as precursors to other amino acids, metabolites or biologically active molecules, or have their own bioactivity.

The metabolism of amino acids and the bioactivity of amino acid metabolites is summed up in Table A2.

Some other important facts about the role of free amino acids in organisms are summarised in Table A3; however, an exhaustive description of the topic is beyond the scope of this chapter. The reader is referred to the literature and biochemical textbooks [18, 32, 33].

Table A1 Biosynthesis of α-Amino acids. (→ … → denotes "more than one step")

AA	Synthesized from	Biochemical cycle	Interrelated with	Ref.
Alanine	Glutamate and pyruvate			[A1]
Arginine	glutamate →…→ ornithine →…→ citruline →…→	Urea cycle		[A2]
Asparagine	Aspartate and ammonia			[A3]
	Aspartate and glutamate			
Aspartic acid	Glutamate and oxaloacetate			[A3]
Cysteine	homocysteine and serine → cystathione			[A4–A7]
Glutamine	Glutamate an and ammonia			[A8]
Glutamic acid	α-ketoglutarate and ammonia		Arginine, glutamine, proline	[A9]
	α-ketoglutarate and an α-amino acid			
	Glutamine			
Glycine	Serine and tetrahydrofolate	Active methyl cycle		[A9]
Histidine	Histidinol	Metabolism of pentoses and purines		[A9]
Isoleucine	Threonine →…→ 2-oxo-3-methylvalerate		Leucine, valine	[A10]
Leucine	threonine →…→ α-ketoisovalerate → … → 4-methyl-2-oxopentanoate		Isoleucine, valine	[A10, A11]
Lysine	Aspartate →…→ meso-2,6-diaminoheptanedionate		Threonine	[A12]
Methionine	Aspartate →…→ homocysteine	Active methyl cycle		[A13]
Phenylalanine	Phenylpyruvate	Prephenate branch of shikamate pathway	Tyrosine	[A14, A15]
	Pretyrosine			
Proline	Glutamate → … → δ^1-pyrroline-5-carboxylate			[A16]
	In plants: ornithine			
Serine	3-phosphoglycerate → … → 3-phosphoserine			[A17]
Threonine	Aspartate →…→ homoserine		Lysine	[A3, A18]
Tryptophan	Indole and serine	Anthranilate branch of shikamate pathway		[A14, A19]
Tyrosine	Phenylalanine	Prephenate branch of shikamate pathway	Phenylalanine	[A14, A20]
	p-hydroxyphenylpyruvate			
	Pretyrosine			
Valine	Threonine → … → 2-oxoisovalerate		Isoleucine, leucine	[A21]
Hydroxyproline	Proline (in bacteria)			[A22, A23]
	Proline residues in collagen (in plants and animals)			

Table A2 Metabolism of amino acids

AA	Skeleton catabolism	Biosynthesis of the other AA	Other metabolites
Alanine	Glucogenic		Pyruvate → … → glucose [A24]
Arginine	Glucogenic	Ornithine [A30]	Nitric oxide [A29, A30]
			Urea [A30]
			Polyamines [A30]
			Creatine [A30]
			Agmatine (a novel neurotransmitter in brain involved in learning and memory and exhibiting anticonvulsant, antinociceptive, anxiolytic and antidepressant-like actions. [A32]
Asparagine	Glucogenic	Aspartate [A3]	
Aspartic acid	Glucogenic	Asparagine homoserine, threonine, isoleucine, methionine, lysine [A3, A10, A11, A18, A21]	Inosine (the precursor to the purine bases) [A9]
Cysteine	Glucogenic		Taurine (essential for normal brain and skeletal muscle functioning) [A33]
Glutamine	Glucogenic	Glutamate [A8]	Nitrogen donation for the synthesis of purines [A9]
Glutamic acid	Glucogenic	Glutamine arginine, proline [A8, A2, A16]	GABA (inhibitory neurotransmitter) [A27]
Glycine	Glucogenic		Donation of C_2N subunit for the synthesis of purines [A9]
			d-aminolevulinic acid (the key precursor to porphyrins) [A9]
Histidine	Glucogenic		Trans-urocanic acid → glutamic acid (in UV light, the urocanic acid undergos trans - cis isomerisation and both isomers, of absorption spectra overlapping to a large extent those of nucleic acids, forms a natural sun screen for the skin.) [A34]
			Carnosine (pH-buffering role in muscles for overcoming muscle fatigue, antioxidant properties, metal-chelator) [A35]

Table A2 (continued)

AA	Skeleton catabolism	Biosynthesis of the other AA	Other metabolites
			Histamine (gastric acid secretion and smoothing muscle relaxation and inhibiting of neurotransmitter release) [A28]
Isoleucine	Glucogenic and ketogenic		In yeasts: amyl alcohol [A36]
Lleucine	Ketogenic		β-hydroxy-β-methylbutyrate (involved in cholesterol biosynthesis) [A37, A38] In yeasts: isoamyl alcohol [A36]
Lysine	Ketogenic		Cadaverine (responsible for distinctive odors of urine and semen as well as of corpse; in plants it is essential for the normal root development) [A40] Saccharopine [A38] Carnitine (transport of fatty acids from the cytosol into the mitochondria) [A41]
Methionine	Glucogenic	Cysteine, homocysteine [A4, A5–A7]	S-Adenosylmethionine (methyl donor for tens of transmethylation reactions of nucleic acids, proteins, lipids, and secondary metabolites) [A42] Carnitine (transport of fatty acids from the cytosol into the mitochondria) [A41] Ethylene (the plant hormone) [A42]
Phenylalanine	Glucogenic and ketogenic	Tyrosine [A14, A20]	Monolignols → lignin (confers structural support, vascular integrity, and pathogen resistance to plants) [A92, A93] A variety of phenylpropanoid compounds generated at the phenylpropanoid pathway were shown to play miscellaneous roles including increase of the plant disease resistance [A43, A44, A45] Phenylacetic acid (a plant hormone), [A46, A47] Phenylethylamine (a neurotransmitter) [A46, A47] 2-phenylacetamide [A46, A47] Phenylpyruvate [A46, A47] Various aromatic volatiles such as benzaldehyde, benzyl alcohol, phenylacetaldehyde, phenethyl alcohol, and ethyl phenylacetate acetate contributing to plant aroma. [A46, A47]

Table A2 (continued)

AA	Skeleton catabolism	Biosynthesis of the other AA	Other metabolites
Proline	Glucogenic	Hydroxyproline [A22, A23]	
Serine	Glucogenic	Cysteine, glycine, tryptophan [A4, A5–A7, A9, A14, A19]	Synthesis of sphingolipids (signaling molecules involved in various cellular activities). [A48] Phosphatidylserine (component of cell membranes) [A9]
Threonine	Glucogenic and ketogenic	Isoleucine, leucine, valine [A10, A11, A21]	
Tryptophan	Glucogenic and ketogenic		Serotonin, i.e. 5-hydroxytryptamine (aminergic neurotransmitter) [A49] Tryptamine (serotonin releasing agent) [A49] Melatonin (neurohormone) [A49] Niacin (vitamin PP) [A49] Camalexin (toxin) [A49, A50] Indole-3-acetic acid (a plant hormone). [A49, A51]
Tyrosine	Glucogenic and ketogenic		Catecholamine neurotransmitters: l-DOPA (dihydroxyphenylalanine), norepinephrine, and epinephrine. [A26] Triiodothyronine (T3) and thyroxine (T4) [A52] Precursor to morphine [A53] Precursor to melanin [A26]
Valine	Glucogenic		In yeasts: isobutyl alcohol [A21]

Table A3 Roles of free amino acids in organisms

AA	Role
Alanine	Glucose—alanine cycle; ammonia carrier [A24]
Arginine	Healing of injuries and damaged tissue [A66, A67]
	Increases growth hormone secretion and thereby influences immune function [A68]
	Supplementation or intravenous administration of arginine is beneficial in improving reproductive, cardiovascular, pulmonary, renal, gastrointestinal, liver and immune functions, as well as enhancing insulin sensitivity [A69]
Asparagine	Presence of free asparagine in blood is an essential factor for the growth of leukaemic blasts [A70–A73]
	In plants: nitrogen storage and nitrogen transport compound [A74]
Aspartic acid	Neurotransmitter: stimulates NMDA receptors [A75]
	Malate-aspartate shuttle [A76]
Cysteine	Precursor to the antioxidant glutathione: playing a critical role in cellular defenses against oxidative stress, free radicals, detoxifyication. [A77, A78]
	Involved in cobalt ion complexation in plant cells [A79]
Glutamine	Nontoxic transporter of ammonia in the blood circulation [A80]
	Regulation of acid-base balance in the kidney by producing ammonium [A80]
	Anabolic action after oral and intravenous administration; reducing of healing times after operations [A81]
Glutamic acid	An excitatory neurotransmitter (while its metabolite GABA is a inhibitory one) [A27]
	Disposal of excess or waste nitrogen [A27]
Glycine	Inhibitory neurotransmitter in the central nervous system [A82]
	There is some evidence of glycine improving sleep quality [A83]
Histidine	In plants the free L-histidine content increases tolerance to Ni and also elevates tolerance to Co and Zn [A84, A85]
Isoleucine	Shuttle molecule for the exchange of nitrogen in the glutamate/glutamine cycle. [A86] disorders in its utilization usually characterized by neuropathological symptoms [A86]
	In plants: stress tolerance [A10]
Leucine	Regulates glucose-alanine cycle [A24]
	l-leucine supplementation during feeding improves muscle protein synthesis in the elderly independently of an overall increase of other amino acids. [A77]
	Effective at attenuating muscle atrophy during exercise stress, and in models of cancer, congestive heart failure, sepsis, and HIV. [A78]
	Shuttle molecule for the exchange of nitrogen in the glutamate/glutamine cycle. [A86] disorders in its utilization usually characterized by neuropathological symptoms [A86]
	In plants: stress tolerance [A10]
Lysine	Inhibits normal replication of herpes simplex virus, accelerates recovery from the infection and suppressing recurrence [A89, A90]
	Serotonin antagonist [A91]
	In plants: related to seed, root elongation and floral development as well as responce to abiotic stress [A39]
Methionine	In the synthesis of eukaryotic proteins methionine is the initiating amino acid usually removed after the protein is synthesized [A92]

Table A3 (continued)

AA	Role
	Regulation of the balance between methionine and cysteine for protein synthesis [A4]
Phenylalanine	
Proline	Proline oxidase exhibits the tumor-suppressive effects [A93]
	In plants: accumulated in response to various stresses (drought, high temperature, low temperature etc.), it confers osmotic tolerance during stress conditions [A94]
Serine	
Threonine	Critically important for stem cells development [A95]
Tryptophan	
Tyrosine	Both meta- and ortho- isomers of l-tyrosine inhibited tumor growth [A96]
Valine	l-valine deficient diet may be involved in the central mechanism of anorexia [A97]
	Shuttle molecules for the exchange of nitrogen in the glutamate/glutamine cycle. [A86] disorders in its utilization usually characterized by neuropathological symptoms [A86]
	In plants: stress tolerance [A10]
Hydroxyproline	Scavenges oxidants and regulates the redox state of cells [A98]

Appendix References

[A1] Felig P (1973) Glucose-alanine cycle. Metab Clin Exper 22:179–207
[A2] Wu G, Morris Jr SM (1998) Arginine metabolism: nitric oxide and beyond. Biochem J 336:1–17
[A3] Azevedo RA, Lancien M, Lea PJ (2006) The aspartic acid metabolic pathway, an exciting and essential pathway in plants. Amino Acids 30:143–162
[A4] MacCoss MJ, Fukagawa NK, Matthews DE (2001) Measurement of intracellular sulfur amino acid metabolism in humans. Am J Physiol Endocrinol Metab 280:E947–E955
[A5] Williams RA, Westrop GD, Coombs GH (2009) Two pathways for cysteine biosynthesis in Leishmania major. Biochem J 420:451–462
[A6] Muller S, Liebau E, Walter RD, Krauth-Siegel RL (2003) Thiol based redox metabolism of protozoan parasites. Trends Parasitol 19:320–328
[A7] Morzycka E, Paszewski A (1979) Two pathways of cysteine biosynthesis in Sacharomycopsis lipolytica. FEBS Lett 101:97–100
[A8] Forde BG, Lea PJ (2007) Glutamate in plants: metabolism, regulation, and signalling. J Exp Bot 58:2339–2358
[A9] Moat AG, Foster JW, Spector MP (2002) Biosynthesis and metabolism of amino acids. In: Microbial Physiology, Wiley-Liss, Inc, New York, pp 503–545
[A10] Joshi V, Joung JG, Fei Z, Jander G (2010) Interdependence of threonine, methionine and isoleucine metabolism in plants: accumulation and transcriptional regulation under abiotic stress. Amino Acids 39:933–947
[A11] Kohlhaw GB (2003) Leucine biosynthesis in fungi: entering metabolism through the back door. Microbiol Mol Biol Rev 67:1–15
[A12] Galili G (1995) Regulation of lysine and threonine synthesis. Plant Cell 7:899–906
[A13] Brosnan JT, Brosnan ME, Bertolo RFP, Brunton JA (2007) Methionine: a metabolically unique amino acid. Livestock Science 112:2–7
[A14] Herrmann KM (1995) The Shikimate pathway: early steps in the biosynthesis of aromatic compounds. Plant Cell 7:907–919

[A15] Maeda H, Yoo H, Dudareva N (2011) Prephenate aminotransferase directs plant phenylalanine biosynthesis via arogenate. Nat Chem Biol 7:19–21
[A16] Kavi Kishor PB, Sangam S, Amrutha RN, Sri Laxmi P, Naidu KR, Rao KRSS, Rao S, Reddy KJ, Theriappan P, Sreenivasulu N (2005) Regulation of proline biosynthesis, degradation, uptake and transport in higher plants: its implications in plant growth and abiotic stress tolerance. Curr Sci 88:424–438
[A17] Ho CL, Saito K (2001) Molecular biology of the plastidic phosphorylated serine biosynthetic pathway in Arabidopsis thaliana. Amino Acids 20:243–259
[A18] Willis LB, Lessard PA, Sinskey AJ (2005) Synthesis of L-threonine and branched-chain amino acids. In: Eggeling L, Bott M (eds) Handbook of Corynebacterium glutamicum. CRC Press Taylor & Francis Group, Boca Raton, pp 511–526
[A19] Gu P, Yang F, Kang J, Wang Q, Qi Q (2012) One-step of tryptophan attenuator inactivation and promoter swapping to improve the production of L-tryptophan in Escherichia coli. Microb Cell Fact 11:1–9
[A20] Fronza G, Fuganti C, Schmidt HL, Werner RA (2002) The δ18O-value of the p-OH group of L-tyrosine permits the assignment of its origin to plant or animal sources. Eur Food Res Technol 215:55–58
[A21] Dickinson JR, Harrison SJ, Hewlins MJ (1998) An investigation of the metabolism of valine to isobutyl alcohol in Saccharomyces cerevisiae. J Biol Chem 273:25751–25756
[A22] Katz E, Kamal F, Mason K (1979) Biosynthesis of trans-4-hydroxy-L-proline by Streptomyces griseoviridus. J Biol Chem 254:6684–6690
[A23] White CE, Gavina JMA, Morton R, Britz-McKibbin P, Finan TM (2012) Control of hydroxyproline catabolism in Sinorhizobium meliloti. Mol Microbiol 85:1133–1147
[A24] Felig P (1973) The glucose-alanine cycle. Metabolism 22:179–207
[A25] D'Mello JPF (2003) Amino acids as multifunctional molecules. In: D'Mello JPF (ed) Amino acids in animal nutrition. CABI Publishing, Wallingford, pp 1–14
[A26] Daubner SC, Le T, Wang S (2011) Tyrosine hydroxylase and regulation of dopamine synthesis. Arch Biochem Biophys 508:1–12
[A27] Bak LK, Schousboe A, Waagepetersen HS (2006) The glutamate/GABA-glutamine cycle: aspects of transport, neurotransmitter homeostasis and ammonia transfer. J Neurochem 98:641–653
[A28] Hill SJ, Ganellin CR, Timmerman H, Schwartz JC, Shankley NP, Young JM, Schunack W, Levi R, Haas HL (1997) International Union of Pharmacology. XIII. Classification of histamine receptors. Pharmacol Rev 49:253–278
[A29] Moncada S, Higgs A (1993) The L-arginine-nitric oxide pathway. New Eng J Med 329:2002–2012
[A30] Wu G, Morris Jr SM (1998) Arginine metabolism: nitric oxide and beyond. Biochem J 336:1–17
[A31] Brosnan JT, Brosnan ME, Bertolo RFP, Brunton JA (2007) Methionine: a metabolically unique amino acid. Livestock Science 112:2–7
[A32] Uzbay TI (2012) The pharmacological importance of agmatine in the brain. Neurosci Biobehav Rev 36:502–519
[A33] Hayes KC (1976) A review on the biological function of taurine. Nutr Rev 34:161–165
[A34] Barbatti M (2011) The role of tautomers in the UV absorption of urocanic acid. Phys Chem Chem Phys 13:4686–4692
[A35] Gariballa SE, Sinclair AJ (2000) Carnosine: physiological properties and therapeutic potential. Age Ageing 29:207–210
[A36] Dickinson JR, Harrison SJ, Dickinson JA, Hewlins MJ (2000) An investigation of the metabolism of isoleucine to active Amyl alcohol in Saccharomyces cerevisiae. J Biol Chem 275:10937–10942
[A37] Nissen SL, Abumrad NN (1997) Nutritional role of the leucine metabolite β-hydroxy β-methylbutyrate (HMB). J Nutr Biochem 8:300–311
[A38] Pimentel GD, Rosa JC, Lira FS, Zanchi NE, Ropelle ER, Oyama LM, Oller do Nascimento CM, de Mello MT, Tufik S, Santos RVT (2011) β-Hydroxy-β-methylbutyrate (HMβ)

supplementation stimulates skeletal muscle hypertrophy in rats via the mTOR pathway. Nutr Metab (Lond) 8:1–7
[A39] Arruda P, Kemper EL, Papes F, Leite A (2000) Regulation of lysine catabolism in higher plants. Trends Plant Sci 5:324–330
[A40] Gamarnik A, Frydman RB (1991) Cadaverine, an essential diamine for the normal root development of germinating soybean (Glycine max) seeds. Plant Physiol 97:778–785
[A41] Rebouche CJ (2012) L-Carnitine, In: Erdman JW, Macdonald IA, Zeisel SH (eds) Present knowledge in nutrition. Wiley-Blackwell, pp 391–404
[A42] Shelly CL (2000) S-Adenosylmethionine. Int J Bioch Cell Biol 32:391–395
[A43] Weng JK, Chapple C (2010) The origin and evolution of lignin biosynthesis. New Phytol 87:273–285
[A44] Fraser CM, Chapple C (2011) The phenylpropanoid pathway in Arabidopsis. Arabidopsis Book 9:e0152–0171
[A45] Solecka D (1997) Role of phenylpropanoid compounds in plant responses to different stress factors. Acta Physiologiae Plantarum 19:257–268
[A46] Yudistira H, McClarty L, Bloodworth RA, Hammond SA, Butcher H, Mark BL, Cardona ST (2011) Phenylalanine induces Burkholderia cenocepacia phenylacetic acid catabolism through degradation to phenylacetyl-CoA in synthetic cystic fibrosis sputum medium. Microb Pathog 51:186–193
[A47] Gonda I, Bar E, Portnoy V, Lev S, Burger J, Schaffer AA, Tadmor Y, Gepstein S, Giovannoni JJ, Katzir N, Lewinsohn E (2010) Branched-chain and aromatic amino acid catabolism into aroma volatiles in Cucumis melo L. fruit. J Exp Bot 61:1111–1123
[A48] Teng C, Dong H, Shi L, Deng Y, Mu J, Zhang J, Yang X, Zuo J (2008) Serine Palmitoyltransferase, a key enzyme for de novo synthesis of sphingolipids, is essential for male gametophyte development in Arabidopsis. Plant Physiol 146:1322–1332
[A49] Sainio EL, Pulkki K, Young SN (1996) L-tryptophan: biochemical, nutritional and pharmacological aspects. Amino Acids 10:21–47
[A50] Glawischnig E (2007) Camalexin. Phytochem 68:401–406
[A51] Won C, Shen X, Mashiguchi K, Zheng Z, Dai X, Cheng Y, Kasahara H, Kamiya Y, Chory J, Zhao Y (2011) Conversion of tryptophan to indole-3-acetic acid by tryptophan aminotransferases of Arabidopsis and YUCCAs in Arabidopsis. Proc Natl Acad Sci USA 108:18518–18523
[A52] Sanjay RC, Kris CD, Rajarathnam ER (2010) Chemistry of thyroxine: an historical perspective and recent progress on its synthesis. Tetrahedron 66:1955–1962
[A53] Poeaknapo C, Schmidt J, Brandsch M, Dräger B, Zenk MH (2004) Endogenous formation of morphine in human cells. Proc Natl Acad Sci USA 101:14091–14096
[A54] Walsh CT (1989) Enzymes in the D-alanine branch of bacterial cell wall peptidoglycan assembly. J Biol Chem 264:2393–2396
[A55] Ghosh AS, Chowdhury C, Nelson DE (2008) Physiological functions of D-alanine carboxypeptidases in Escherichia coli. Trends Microbiol 16:309–317
[A56] Abe H, Yoshikawa N, Sarower MG, Okada S (2005) Physiological function and metabolism of free D-alanine in aquatic animals. Biol Pharm Bull 28:1571–1577
[A57] Morikawa A, Hamase K, Zaitsu K (2003) Determination of D-alanine in the rat central nervous system and periphery using column-switching high-performance liquid chromatography. Anal Biochem 312:66–72
[A58] Fisher GH, D'Aniello A, Vetere A, Padula L, Cusano G, Man EH (1991) Free D-aspartate and D-alanine in normal and Alzheimer brain. Brain Res Bull 26:983–985
[A59] D'Aniello A, Vetere A, Fisher GH, Cusano G, Chavez M, Petrucelli L (1992) Presence of D-alanine in proteins of normal and Alzheimer human brain. Brain Res 592:44–48
[A60] Nagata Y, Masui R, Akino T (1992) The presence of free D-serine, D-alanine, and D-proline in human plasma. Experientia 48:986–988
[A61] Hamase K, Konno R, Morikawa A, Zaitsu K. (2005) Sensitive determination of D-amino acids in mammals and the effect of D-amino-acid oxidase activity on their amounts. Biol Pharm Bull 28:1578–1584

[A62] D'Aniello A (2007) D-Aspartic acid: an endogenous amino acid with an important neuroendocrine role. Brain Res Rev 53:215–234
[A63] D'Aniello S, Somorjai I, Garcia-Fernàndez J, Topo E, D'Aniello A (2011) D-Aspartic acid is a novel endogenous neurotransmitter. FASEB J 25:1014–1027
[A64] D'Aniello S, Spinelli P, Ferrandino G, Peterson K, Tsesarskaja M, Fisher GH, D'Aniello A (2005) Cephalopod vision involves dicarboxylic amino acids: D-aspartate, L-aspartate and L-glutamate. Biochem J 386:331–340
[A65] Schell MJ, Molliver ME, Snyder SH (1995) D-serine, an endogenous synaptic modulator: localization to astrocytes and glutamate-stimulated release. Proc Natl Acad Sci U S A 92:3948–3952
[A66] Stechmiller JK (2005) Arginine supplementation and wound healing. Nutr Clin Pract 20:52–61
[A67] Witte MB, Barbul A (2003) Arginine physiology and its implication for wound healing. Wound Rep Regen 11:419–423
[A68] Daly JM, Reynolds J, Thom A, Kinsley L, Dietrick-Gallagher M, Shou J, Ruggieri B (1988) Immune and metabolic effects of arginine in the surgical patient. 208:512–523
[A69] Wu G, Bazer FW, Davis TA, Kim SW, Li P, Rhoads JM, Satterfield MC, Smith SB, Spencer TE, Yin Y (2009) Arginine metabolism and nutrition in growth, health and disease. Amino Acids 37:153–168
[A70] Broome JD (1968) Studies on the mechanism of tumor inhibition by L-asparaginase. Effects of the enzyme on asparagine levels in the blood, normal tissues, and 6C3HED lymphomas of mice: differences in asparagine formation and utilization in asparaginase-sensitive and -resistant lymphoma cells. J Exp Med 127:1055–1072
[A71] Boos J, Werber G, Ahlke E, Schulze-Westhoff P, Nowak-Göttl U, Würthwein G, Verspohl EJ, Ritter J, Jürgens H (1996) Monitoring of asparaginase activity and asparagine levels in children on different asparaginase preparations. Eur J Cancer 32A:1544–1550
[A72] Chakrabarti R, Schuster SM (1997) L-asparaginase. Perspectives on the mechanisms of action and resistance. Int J Pediatr Hematol/Oncol 4:597–611
[A73] Avramis VI, Panosyan EH (2005) Pharmacokinetic/pharmacodynamic relationships of asparaginase formulations: the past, the present and recommendations for the future. Clin Pharmacokinet 44:367–393
[A74] Lea PJ, Sodek L, Parry MAJ, Shewry PR, Halford NG (2007) Asparagine in plants. Ann Appl Biol 150:1–26
[A75] Stone TW, Burton NR (1988) NMDA receptors and ligands in the vertebrate CNS. Prog Neurobiol 30:333–368
[A76] Barron JT, Gu L, Parrillo JE (1988) Malate-aspartate shuttle, cytoplasmic NADH redox potential, and energetics in vascular smooth muscle. J Mol Cell Cardiol 30:1571–1579
[A77] Griffith OW (1999) Biologic and pharmacologic regulation of mammalian glutathione synthesis. Free Radic Biol Med 27:922–935
[A78] Lo M, Wang YZ, Gout PW (2008) The x(c)-cystine/glutamate antiporter: a potential target for therapy of cancer and other diseases. J Cell Physiol 215(3):593–602
[A79] Oven M, Grill E, Golan-Goldhirsh A, Kutchan TM, Zenk MH (2002) Increase of free cysteine and citric acid in plant cells exposed to cobalt ions. Phytochemistry 60:467–474
[A80] Lacey JM, Wilmore DW (1990) Is glutamine a conditionally essential amino acid? Nutr Rev 48:297–309
[A81] Curi R, Lagranha CJ, Doi SQ, Sellitti DF, Procopio J, Pithon-Curi TC, Corless M, Newsholme P (2005) Molecular mechanisms of glutamine action. J Cell Physiol 204:392–401
[A82] Carmans S, Hendriks JJ, Thewissen K, Van den Eynden J, Stinissen P, Rigo JM, Hellings N (2010) The inhibitory neurotransmitter glycine modulates macrophage activity by activation of neutral amino acid transporters. J Neurosci Res 88:2420–2430
[A83] Yamadera W, Inagawa K, Chiba S, Bannai M, Takahashi M, Nakayama K (2007) Glycine ingestion improves subjective sleep quality in human volunteers, correlating with polysomnographic changes. Sleep Biol Rhyt 5:126–131

[A84] Rees JD, Ingle RA, Smith JA (2009) Relative contributions of nine genes in the pathway of histidine biosynthesis to the control of free histidine concentrations in Arabidopsis thaliana. Plant Biotechnol J 7:499–511
[A85] Verbruggen N, Hermans C, Schat H (2009) Molecular mechanisms of metal hyperaccumulation in plants. New Phytologist 181:759–776
[A86] Murín R, Mohammadi G, Leibfritz D, Hamprecht B (2009) Glial metabolism of valine. Neurochem Res 34:1195–1203
[A87] Rieu I, Balage M, Sornet C, Giraudet C, Pujos E, Grizard J, Mosoni L, Dardevet D. (2006) Leucine supplementation improves muscle protein synthesis in elderly men independently of hyperaminoacidaemia. J Physiol 575:305–315
[A88] Pimentel GD, Rosa JC, Lira FS, Zanchi NE, Ropelle ER, Oyama LM, Oller do Nascimento CM, de Mello MT, Tufik S, Santos RVT (2011) β-Hydroxy-β-methylbutyrate (HMβ) supplementation stimulates skeletal muscle hypertrophy in rats via the mTOR pathway. Nutr Metab (Lond) 8:1–7
[A89] Singh M, Muralidhara Rao D, Pande S, Battu S, Mahalakshmi K, Rajeswar Dutt K, Ramesh M (2011) Medicinal uses of L-Lysine: past and future. Int J Res Pharm Sci 2:637–642
[A90] Singh BB, Udani J, Vinjamury SP, Der-Martirosian C, Gandhi S, Khorsan R, Nanjegowda D, Singh V (2005) Safety and effectiveness of an L-lysine, zinc, and herbal-based product on the treatment of facial and circumoral herpes. Altern Med Rev 10:123–127
[A91] Smriga M, Torii K (2003) L-Lysine acts like a partial serotonin receptor 4 antagonist and inhibits serotonin-mediated intestinal pathologies and anxiety in rats. Proc Natl Acad Sci USA 100:15370–15375
[A92] Drabkin HJ, RajBhandary UL (1998) Initiation of protein synthesis with codons other than AUG and amino acids other than methionine. Mol Cell Biol 15:5140–5147
[A93] Tanner JJ (2008) Structural biology of proline catabolism. Amino Acids 35:719–730
[A94] Verbruggen N, Hermans C (2008) Proline accumulation in plants: a review. Amino Acids 35:753–759
[A95] Wang J, Alexander P, Wu L, Hammer R, Cleaver O, McKnight SL (2009) Dependence of mouse embryonic stem cells on threonine catabolism. Science 325:435–439
[A96] Ruggiero RA, Bruzzo J, Chiarella P, Bustuoabad OD, Meiss RP, Pasqualini CD (2012) Concomitant tumor resistance: the role of tyrosine isomers in the mechanisms of metastases control. Cancer Res 72:1043–1050
[A97] Nakahara K, Takata S, Ishii A, Nagao K, Bannai M, Takahashi M, Murakami N (2012) Somatostatin is involved in anorexia in mice fed a valine-deficient diet. Amino Acids 42:1397–1404
[A98] Wu G, Bazer FW, Burghardt RC, Johnson GA, Kim SW, Knabe DA, Li P, Li X, McKnight JR, Satterfield MC, Spencer TE (2011) Proline and hydroxyproline metabolism: implications for animal and human nutrition. Amino Acids 40:1053–1063
[A99] Wendisch VF (ed) (2007) Amino acid biosynthesis—pathways, regulation and metabolic engineering. Springer
[A100] Bender DA (2012) Amino acid metabolism. Wiley-Blackwell

References

1. Feynman RP, Leighton RB, Sands M (1964) The Feynman lectures on Physics, vol I. Addison–Wesley, USA, pp 3–10 “A poet once said “The whole universe is in a glass of wine.” We will probably never know in what sense he meant that, for poets do not write to be understood. But it is true that if we look at a glass of wine closely enough we see the entire universe. There are the things of physics: the twisting liquid which evaporates depending on the wind and weather, the reflections in the glass, and our imagination adds the atoms. The glass is a distillation of the Earth’s rocks, and in its composition we see the secrets of the universe’s age, and the evolution of stars. What strange arrays of chemicals are in the wine? How did

they come to be? There are the ferments, the enzymes, the substrates, and the products. There in wine is found the great generalization: all life is fermentation. Nobody can discover the chemistry of wine without discovering, as did Louis Pasteur, the cause of much disease. How vivid is the claret, pressing its existence into the consciousness that watches it! If our small minds, for some convenience, divide this glass of wine, this universe, into parts—physics, biology, geology, astronomy, psychology, and so on—remember that Nature does not know it! So let us put it all back together, not forgetting ultimately what it is for. Let it give us one more final pleasure: drink it and forget it all!"

2. Jamróz MH, Rode JE, Ostrowski S, Lipiński PFJ, Dobrowolski JCz (2012) Chirality measures of α-amino acids. J Chem Inform Model 52:1462–1479
3. Van Hall G, Saltin B, Van der Vusse GJ, Soderlund K, Wagenmakers AJM (1995) Deamination of amino acids as a source for ammonia production in human skeletal muscle during prolonged exercise. J Physiol 489:251–261
4. Felig P (1973) The glucose-alanine cycle. Metabolism 22:179–207
5. Perriello G, Jorde R, Nurjhan N, Stumvoll M, Dailey G, Jenssen T, Bier DM, Gerich JE (1995) Estimation of glucose-alanine-lactate-glutamine cycles in postabsorptive humans: role of skeletal muscle. Am J Physiol 269:E443–E450
6. D'Mello JPF (2003) Amino acids as multifunctional molecules. In: D'Mello JPF (ed) Amino acids in animal nutrition. CABI Publishing, Wallingford, pp 1–14
7. Sauer U, Eikmanns BJ (2005) The PEP-pyruvate-oxaloacetate node as the switch point for carbon flux distribution in bacteria FEMS. Microbiol Rev 29:765–794
8. Daubner SC, Le T, Wang S (2011) Tyrosine hydroxylase and regulation of dopamine synthesis. Arch Biochem Biophys 508:1–12
9. Bak LK, Schousboe A, Waagepetersen HS (2006) The glutamate/GABA-glutamine cycle: aspects of transport, neurotransmitter homeostasis and ammonia transfer. J Neurochem 98:641–653
10. Hill SJ, Ganellin CR, Timmerman H, Schwartz JC, Shankley NP, Young JM, Schunack W, Levi R, Haas HL (1997) International Union of Pharmacology. XIII. Classification of histamine receptors. Pharmacol Rev 49:253–278
11. Blows WT (2000) Neurotransmitters of the brain: serotonin, noradrenaline (norepinephrine), and dopamine. J Neurosci Nursing 32:234–238
12. Serretti A, Olgiati P (2008) Biochemistry of depressive disorders. Role of serotonin, amino acid neurotransmitters, substance P and neurosteroids. Clinical Neuropsychiatry 5:225–239
13. Moncada S, Higgs A (1993) The L-arginine-nitric oxide pathway. New Eng J Med 329:2002–2020
14. Wu G, Morris SM Jr (1998) Arginine metabolism: nitric oxide and beyond. Biochem J 336:1–17
15. Krajčovičová-Kudláčková M, Šimončič R, Béderová A, Babinská K, Béder I (2000) Correlation of carnitine levels to methionine and lysine intake. Phys Res 49:399–402
16. Shelly CL (2000) S-Adenosylmethionine. Int J Bioch Cell Biol 32:391–395
17. Finkelstein JD, Martin JJ (1984) Methionine metabolism in mammals. Distribution of homocysteine between competing pathways. J Biol Chem 259:9508–9513
18. Moat AG, Foster JW, Spector MP (2002) Biosynthesis and metabolism of amino acids. In: Microbial physiology. Wiley-Liss, Inc, New York, pp 503–545
19. Gibson KD, Laver WG, Neuberger A (1958) Initial stages in the biosynthesis of porphyrins. 2. The formation of delta-aminolaevulic acid from glycine and succinyl-coenzyme A by particles from chicken erythrocytes. Biochem J 70:71–81
20. Walsh CT (1989) Enzymes in the D-Alanine branch of bacterial cell wall peptidoglycan assembly. J Biol Chem 264:2393–2396.
21. Ghosh AS, Chowdhury C, Nelson DE (2008) Physiological functions of D-alanine carboxypeptidases in Escherichia coli. Trends Microbiol 16:309–317
22. Abe H, Yoshikawa N, Sarower MG, Okada S (2005) Physiological function and metabolism of free D-alanine in aquatic animals. Biol Pharm Bull 28:1571–1577
23. Morikawa A, Hamase K, Zaitsu K (2003) Determination of D-alanine in the rat central nervous system and periphery using column-switching high-performance liquid chromatography. Anal Biochem 312:66–72

24. Fisher GH, D'Aniello A, Vetere A, Padula L, Cusano G, Man EH (1991) Free D-aspartate and D-alanine in normal and Alzheimer brain. Brain Res Bull 26:983–985
25. D'Aniello A, Vetere A, Fisher GH, Cusano G, Chavez M, Petrucelli L (1992) Presence of D-alanine in proteins of normal and Alzheimer human brain. Brain Res 592:44–48
26. Nagata Y, Masui R, Akino T (1992) The presence of free D-serine, D-alanine, and D-proline in human plasma. Experientia 48:986–988
27. Hamase K, Konno R, Morikawa A, Zaitsu K (2005) Sensitive determination of D-amino acids in mammals and the effect of D-amino-acid oxidase activity on their amounts. Biol Pharm Bull 28:1578–1584
28. D'Aniello A (2007) D-Aspartic acid: an endogenous amino acid with an important neuroendocrine role. Brain Res Rev 53:215–234
29. D'Aniello S, Somorjai I, Garcia-Fernàndez J, Topo E, D'Aniello A (2011) D-aspartic acid is a novel endogenous neurotransmitter. FASEB J 25:1014–1027
30. D'Aniello S, Spinelli P, Ferrandino G, Peterson K, Tsesarskaja M, Fisher GH, D'Aniello A (2005) Cephalopod vision involves dicarboxylic amino acids: D-aspartate, L-aspartate and L-glutamate. Biochem J 386:331–340
31. Schell MJ, Molliver ME, Snyder SH (1995) D-serine, an endogenous synaptic modulator: localization to astrocytes and glutamate-stimulated release. Proc Natl Acad Sci U S A 92:3948–3952
32. Wendisch VF (ed) (2007) Amino acid biosynthesis—pathways, regulation and metabolic engineering. In: Steinbüchel (ed) Microbiology Monographs, Springer
33. Bender DA (2012) Amino acid metabolism. John Wiley & Sons
34. Belitz H-D, Grosch W, Schieberle P (2009) Amino acids, peptides, proteins. In: Food Chemistry Springer, pp 8–92
35. Mizerski W (2008) Chemical tables (in Polish Tablice Chemiczne) Adamantan, Warsaw
36. Gridchin SN, Romodanovskii PA, Pyreu DF (2009) The heats of interaction of L-glutamine and L-glutamic acid with KOH and HNO_3 in aqueous solutions. Rus J Phys Chem A 83:138–142
37. Kumar NV, Rao GN (2011) Influence of dielectric constant on protonation equilibria of L-Aspartic acid in acetonitrile- and ethylene glycol-water mixtures. Acta Chim Sloven 58:342–346
38. Zhang P, Polavarapu PL (2006) Vibrational circular dichroism of matrix-assisted amino acid films in the mid-infrared region. Appl Spectrosc 60:378–385
39. Ide M, Maeda Y, Kitano H (1997) Effect of hydrophobicity of amino acids on the structure of water. J Phys Chem B 101:7022–7026
40. McNaught AD, Wilkinson A (1997) IUPAC. Compendium of chemical terminology, Blackwell Scientific Publications, Oxford. http://goldbook.iupac.org/
41. Biswas KM, DeVido DR, Dorsey JG (2003) Evaluation of methods for measuring amino acid hydrophobicities and interactions. J Chromatogr A 1000:637–655
42. Levitt M (1976) A simplified representation of protein conformation for rapid simulations of protein folding. J Mol Biol 104:59–107
43. Parker JMR, Guo D, Hodges RS (1986) New hydrophilicity scale derived from high performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. Biochem 25:5425–5432
44. Radzicka A, Wolfenden R (1988) Comparing the polarities of the amino acids: side-chain distribution coefficients between the vapor phase, cyclohexane, 1-octanol, and neutral aqueous solution. Biochem 27:1664–1670
45. Kyte K, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157:105–132
46. Chothia C (1974) Hydrophobic bonding and accessible surface area in proteins. Nature 248:338–339
47. Janin J (1979) Surface and inside volumes in globular proteins. Nature 277:491–492
48. Rose GD, Geselowitz AR, Lesser GJ, Lee RH, Zehfus MH (1985) Hydrophobicity of amino acid residues in globular proteins. Science 229:834–838
49. Moret MA, Zebende GF (2007) Amino acid hydrophobicity and accessible surface area. Phys Rev E 75: 011920–04
50. Vinodhini R, Vijaya MS (2011) Machine learning and sequence labeling techniques for protein secondary structure prediction using hydrophobicity scales. Int J Res Rev Artif Intellig 4:86–90

51. Khafizov K, Staritzbichler R, Stamm M, Forrest LR (2010) A study of the evolution of inverted-topology repeats from LeuT-Fold transporters using AlignMe. Biochem 49:10702–10713.
52. Zhou P, Tian F, Zhang M, Li Z (2006) Applying generalized hydrophobicity scale of amino acids to quantitative prediction of human leukocyte antigen-A*0201-restricted cytotoxic T lymphocyte epitope. Chin Sci Bull 51:1439–1443
53. Nakanishi K, Berova N, Woody RW (eds) (2000) Circular dichroism: principles and Applications, 2nd edn. Wiley, New York
54. Berova N, Polavarapu PL, Nakanishi K, Woody RW (eds) (2012) Comprehensive chiroptical spectroscopy, vol 1 and 2. Wiley, New Jersey
55. Nafie LA (2011) Vibrational optical activity: principles and Applications, Wiley
56. Nafie LA, Dukor RK, Freedman TB (2002) Vibrational circular dichroism. In: Chalmers JM, Griffiths PR (eds) Handbook of vibrational spectroscopy. Wiley, Chichester, pp 731–744
57. Polavarapu PL (2006) Quantum mechanical predictions of chirooptical vibrational properties. Int J Quantum Chem 106:1809–1814
58. Polavarapu PL (2007) Renaissance in chiroptical spectroscopic methods for molecular structure determination. Chem Record 7:125–136
59. Polavarapu PL (2008) Why is it important to simultaneously use more than one chiroptical spectroscopic method for determining the structures of chiral molecules. Chirality 20:664–672
60. Barron LD (2004) Molecular light scattering and optical activity, 2nd edn. Cambridge University Press, Cambridge
61. Autschbach J (2009) Computing chiroptical properties with first-principles theoretical methods: background and illustrative examples. Chirality 21:E116–E152
62. Freedman TB, Cao X, Dukor RK, Nafie LA (2003) Absolute configuration determination of chiral molecules in the solution state using vibrational circular dichroism. Chirality 15:743–758
63. Pecul M, Ruud K (2005) Ab initio calculations of vibrational Raman optical activity. Int J Quantum Chem 104:816–829
64. Petrovic AG, Navarro-Vazquez A, Alonso-Gomze J (2010) From relative to absolute configuration of complex natural products: interplay between NMR, ECD, VCD, and ORD assisted by ab initio calculation. Curr Org Chem 14:1612–1628.
65. Alkorta I, Elguero J (2009) Chirality and chirality recognition. In: Leszczynski J, Shukla MK (eds) Practical aspects of computational chemistry, Springer Science, pp 37–83
66. Yang G, Xu Y (2011) Vibrational circular dichroism spectroscopy of chiral molecules. In: Naaman R, Beratan DN, Waldeck DH (eds) Electronic and magnetic properties of chiral molecules and supramolecular architectures. Springer-Verlag Berlin Heidelberg, Top Curr Chem 298:189–236
67. Sadlej J, Dobrowolski JC, Rode JE (2010) VCD spectroscopy as a novel probe for chirality transfer in molecular interactions. Chem Soc Rev 39:1478–1488
68. Magyarfalvi G, Tarczay G, Vass E (2011) Vibrational circular dichroism, WIREs Comput Mol Sci 1:403–425
69. Buckingham AD (ed) (1994) Vibrational optical activity: from fundamentals to biological applications. Faraday Discuss 99
70. Dudley RJ, Mason F, Peacock RD (1972) Infrared vibrational circular dichroism. J Chem Soc. Chem Commun 1084–1085
71. Chabay I, Hsu EC, Holzwarth G (1972) Infrared circular dichroism measurement between 2000 and 5000 cm^{-1}: PrC3 tartrate complexes. Chem Phys Lett 15:211–214
72. Holzwarth G, Chabay I (1972) Optical activity in molecular vibrations: a coupled oscillator model. J Chem Phys 57:1632–1635
73. Hsu EC, Holzwarth G (1973) Vibrational circular dichroism in crystalline α-$NiSO_4$-$6H_2O$ and α-$ZnSeO_4$-$6H_2O$ between 1900 and 5000 cm^{-1}. J Chem Phys 59:4678–4685
74. Barron LD, Buckingham AD (1975) Rayleigh and Raman optical activity. Annu Rev Phys Chem 26:381–396
75. Buckingham AD, Raab RE (1975) Electric-field-induced differential scattering of right and left circularly polarized light. Proc R Soc Lond Ser A 345:365–377
76. Nafie LA, Keiderling TA, Stephens PJ (1976) Vibrational circular dichroism. J Am Chem Soc 98:2715–2723

77. Nafie LA, Diem M (1979) Theory of high frequency differential interferometry: application to the measurement of infrared circular and linear dichroism via Fourier transform spectroscopy. Appl Spectrosc 33:130–135
78. Schlosser DW, Devlin F, Jalkanen K, Stephens PJ (1982) Vibrational circular dichroism of matrix-isolated molecules. Chem Phys Lett 88:286–291
79. Henderson DO, Polavarapu P (1986) Fourier transform infrared vibrational circular dichroism of matrix-isolated molecules. J Am Chem Soc 108:7110–7111
80. Tarczay G, Magyarfalvi G, Vass E (2006) Towards the determination of the absolute configuration of complex molecular systems: matrix isolation vibrational circular dichroism study of (R)-2-amino-1-propanol. Angew Chem Int Ed 45:1775–1777
81. Pohl G, Perczel A, Vass E, Magyarfalvi G, Tarczay G (2007) A matrix isolation study on Ac-Gly-NHMe and Ac-L-Ala-NHMe, the simplest chiral and achiral building blocks of peptides and proteins. Phys Chem Chem Phys 9:4698–4708
82. Tarczay G, Góbi S, Vass E, Magyarfalvi G (2009) Model peptide–water complexes in Ar matrix: complexation induced conformation change and chirality transfer. Vib Spectrosc 50:21–28
83. Góbi S, Knapp K, Vass E, Majer Z, Magyarfalvi G, Hollósi M, Tarczay G (2010) Is β-homo-proline a pseudo-γ-turn forming element of β-peptides? an IR and VCD spectroscopic study on Ac-β-HPro-NHMe in cryogenic matrices and solutions. Phys Chem Chem Phys 12:13603–13615
84. Polavarapu PL (1998) Vibrational absorption and circular dichroism. In: Vibrational spectra: principles and applications with emphasis on optical activity. Elsevier, New York, pp 143–182
85. Stephens PJ, Devlin FJ, Pan JJ (2008) The determination of the absolute configurations of chiral molecules using VCD spectroscopy. Chirality 20:643–663
86. Helgaker T, Coriani S, Jørgensen P, Kristensen K, Olsen J, Ruud K (2012) Recent advances in wave function-based methods of molecular-property calculations. Chem Rev 112:543–631
87. Stephens PJ, Lowe, MA (1985) Vibrational circular-dichroism. Annu Rev Phys Chem 36:213–241
88. Stephens PJ (1987) Gauge dependence of vibrational magnetic dipole transition moments and rotational strengths. J Phys Chem 91:1712–1715
89. Gaussian 09, Rev. A.1, Frisch MJ, Trucks GW, Schlegel HB et al (2009) Gaussian Inc. Wallingford
90. ADF2009.01, SCM (2009), Theoretical chemistry, Vrije Universiteit, Amsterdam, The Netherlands, http://www.scm.com
91. DALTON, a molecular electronic structure program, Release Dalton2011 (2011), http://daltonprogram.org
92. Amos RD, Alberts IL, Andrews JS, Cohen AJ, Colwell SM, Handy NC, Jayatilaka D, Knowles PJ, Kobayashi R, Laming GJ, Lee AM, Maslen PE, Murray CW, Palmieri P, Rice JE, Simandiras ED, Stone AJ, Su M-D, Tozer DJ, CADPAC 6.5 (1998) The Cambridge analytic derivatives package, Cambridge
93. Bouř P, McCann J, Wieser HJ (1998) The excitation scheme: a new method for calculation of vibrational circular dichroism spectra. J Chem Phys 108:8782–8789
94. Coriani S, Thorvaldsen AJ, Kristensen K, Jorgensen P (2011) Vibrational response-function formulation of vibrational circular dichroism. Phys Chem Chem Phys 13:4224–4229
95. Crawford TD (2006) Ab initio calculation of molecular chiroptical properties. Theor Chem Acc 115:227–245
96. Turney JM, Simmonett AC, Parrish RM, Hohenstein EG, Evangelista F, Fermann JT, Mintz BJ, Burns LA, Wilke JJ, Abrams ML, Russ NJ, Leininger ML, Janssen CL, Seidl ET, Allen WD, Schaefer HF, King RA, Valeev EF, Sherrill CD, Crawford TD (2012) Psi4: an open-source ab initio electronic structure program. WIREs Comput Mol Sci 2:556–565
97. Crawford TD, Sherrill CD, Valeev EF, Fermann JT, King RA, Leininger ML, Brown ST, Janssen CL, Seidl ET, Kenny JP, Allen WD (2007) PSI3: an open-source Ab Initio electronic structure package. J Comput Chem 28:1610–1616
98. Jalkanen KJ, Jürgensen VW, Degtyarenko IM (2005) Linear response properties required to simulate vibrational spectra of biomolecules in various media: (R)-phenyloxirane (A com-

parative theoretical and spectroscopic vibrational study). In: Jensen HJA (ed) Advances in quantum chemistry, vol 50. Response theory and molecular properties, Elsevier Academic Press Inc, San Diego, pp 91–124
99. Atkins PW, Barron LD (1969) Rayleigh scattering of polarized photons by molecules. Mol Phys 16:453–466
100. Barron LD, Buckingham AD (1971) Rayleigh and Raman scattering from optically active molecules. Mol Phys 20:1111–1119
101. Barron LD, Bogaard MP, Buckingham AD (1973) Raman scattering of circularly polarized light by optically active molecules. J Am Chem Soc 95:603–605
102. Hug W, Kint S, Bailey, Scherer JR (1975) Raman circular intensity differential spectroscopy. Spectra of (–)-.alpha.-pinene and (+)-.alpha.-phenylethylamine. J Am Chem Soc 97:5589–5590
103. Nafie LA (2008) Theory of Raman scattering and Raman optical activity: near-resonance theory and levels of approximation. Theo Chem Acc 119:39–55
104. Ruud K, Thorvaldsen AJ (2010) Theoretical approaches to the calculation of Raman optical activity spectra. Chirality 21:E54–E67
105. Cheeseman JR, Frisch MJ (2011) Basis set dependence of vibrational Raman and Raman optical activity intensities, J Comp Theor Chem 7:3323–3334
106. Polavarapu PL (1990) Ab initio Raman and Raman optical activity spectra. J Phys Chem 94:8106–8112
107. Helgaker T, Ruud K, Bak KL, Jørgensen P, Olsen J (1994) Vibrational Raman optical activity calculations using London atomic orbitals. Faraday Discuss 99:165–180
108. Ruud K, Helgaker T, Bouř P (2002) Gauge-origin independent density-functional theory calculations of vibrational Raman optical activity. J Phys Chem A 106:7448–7455
109. Barron LD, Buckingham AD (2010) Vibrational optical activity. Chem Phys Lett 492:199–213
110. Weymuth T, Haag MP, Kiewisch K, Luber S, Schenk S, Jacob CR, Herrmann C, Neugebauer J, Reiher M (2012) MOVIPAC: vibrational spectroscopy with a robust meta-program for massively parallel standard and inverse calculations. J Comput Chem 33:2186–2198
111. Cappelli C, Bloino J, Lipparini F, Barone V (2012) Toward ab Initio anharmonic vibrational circular dichroism spectra in the condensed phase. J Phys Chem Lett 3:1766–1773
112. Bloino J, Biczysko M, Barone V (2012) General perturbative approach for spectroscopy, thermodynamics, and kinetics: methodological background and benchmark Studies. J Chem Theory Comput 8:1015–1036
113. Bloino J, Barone V (2012) A second-order perturbation theory route to vibrational averages and transition properties of molecules: general formulation and application to infrared and vibrational circular dichroism spectroscopies. J Chem Phys 136:124108–124123
114. Jacob CR, Reiher M (2009) Localizing normal modes in large molecules. J Chem Phys 130:084106–084121
115. Tomasi J, Mennucci B, Cammi R (2005) Quantum mechanical continuum solvation models. Chem Rev 105:2999–3093
116. Cramer CJ, Truhlar DG (1999) Implicit solvation models: equilibria, structure, spectra and dynamics. Chem Rev 99:2161–2200
117. Cammi R, Mennucci B, Tomassi J (2003) Computational modelling of the solvent effects on molecular properties: an overview of the polarizable continuum model (PCM) approach. In: Leszczynski J (ed) Computational chemistry, review of current trends, vol 8. World Scientific, Singapore, pp 1–80
118. Rivail JL, Rinaldi D (1996) Liquid-state quantum chemistry: computational applications of the polarizable continuum models. In: Leszczynski J (ed) Computational chemistry, review of current trends, vol 1. World Scientific, Singapore, pp 139–174
119. Monard G, Rivail JL (2012) Solvent effects in quantum chemistry. In Leszczynski J (ed.) Handbook of computational chemistry, Springer, Netherlands, pp. 561–571
120. Pecul M, Ruud K (2007) Solvent effects on natural optical activity. In Mennucci B, Cammi R (eds) Continuum solvation models in chemical physics: from theory to applications. Wiley, VCH, Weinheim, pp 206–219

121. Boys SF, Bernardi F (1970) The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. Mol Phys 19:553–566.
122. Turi L, Dannenberg JJ (1993) Correcting for basis set superposition error in aggregates containing more than two molecules: ambiguities in the calculation of the counterpoise correction. J Phys Chem 97:2488–2490
123. Rode JE, Dobrowolski JCz (2002) Theoretical studies on the oxetane…HCl and thietane…HCl complexes. Chem Phys Lett 360:123–132
124. Tomasi J, Persico M (1994) Molecular interactions in solution: an overview of methods based on continuous distributions of the solvent. Chem Rev 94:2027–2094
125. Cancès E, Mennucci B, Tomasi J (1997) A new integral equation formalism for the polarizable continuum model: theoretical background and applications to isotropic and anisotropic dielectrics. J Chem Phys 107:3032–3041
126. Amovilli C, Mennucci B (1997) Self-consistent-field calculation of Pauli repulsion and dispersion contributions to the solvation free energy in the polarizable continuum Model. J Phys Chem 101:1051–1057
127. Barone V, Cossi M (1998) Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. J Phys Chem A 102:1995–2001
128. Klamt A, Schüürmann G (1993) COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. J Chem Soc Perkin Trans 2:799–805
129. Klamt A (1995) Conductor-like screening model for real solvents: a new approach to the quantitative calculation of solvation phenomena. J Phys Chem 99:2224–2235
130. Klamt A, Jonas V (1996) Treatment of the outlying charge in continuum solvation models. J Chem Phys 105:9972–9981
131. Rivail JL, Rinaldi D (1973) Polarisabilités moléculaires et effet diélectrique de milieu á l'état liquide Étude théorique de la molécule d'eau et de ses diméres. Theor Chim Acta 32:57–70
132. Mikkelsen KV, Dalgaard E, Swanstrøm P (1987) Electron-transfer reactions in solution: an ab initio approach. J Phys Chem 91:3081–3092
133. Mikkelsen KV, Jørgensen P, Jensen HJA (1994) A multiconfiguration self-consistent reaction field response method. J Chem Phys 100:6597–6607
134. Cramer CJ, Truhlar DG (1995) Continuum solvation models: classical and quantum mechanical implementations. In: Lipkowitz KB, Boyd DB (eds) Reviews of computational chemistry, vol 6. VCH, New York, pp 1–72
135. Cramer CJ, Truhlar DG (2006) SMx continuum models for condensed phases. In: Maroulis G, Simos TE (eds) Trends and perspectives in modern computational science. Lecture series on computational methods in sciences and engineering, vol 6. Brill/VSP, Leiden, pp 112–139
136. Cramer CJ, Truhlar DG (2008) A universal approach to solvation modeling. Acc Chem Res 41:760–768
137. Poopari MR, Dezhahang, Yang G, Xu Y (2012) Conformational distributions of N-Acetyl-l-cysteine in aqueous solutions: a combined implicit and explicit solvation treatment of VA and VCD spectra. Chem Phys Chem 13:2310–2321
138. Kaminský J, Šebek J, Bouř P (2009) Molecular dynamics with restrictions derived from optical spectra. J Comp Chem 30:983–991
139. Yang S, Cho M (2009) Direct calculations of vibrational absorption and circular dichroism spectra of alanine dipeptide analogue in water: quantum mechanical/molecular mechanical molecular dynamics simulations. J Chem Phys 131:135102-1–135102-8
140. Kamiński M, Kudelski A, Pecul M (2012) Vibrational optical activity of cysteine in aqueous solution: a comparison of theoretical and experimental spectra. J Phys Chem B 116:4976–4990
141. Hopmann KH, Ruud K, Pecul M, Kudelski A, Dracinsky M, Boŭr P (2011) Explicit versus implicit solvent modeling of Raman optical activity spectra, J Phys Chem B, 115:4128–4137
142. Boŭr P, Keiderling T (2004) Partial optimization of molecular geometry in normal coordinates and use as a tool for simulation of vibrational spectra. J Chem Phys 117:4126–4132

143. Cossi M, Barone V (1998) Analytical second derivatives of the free energy in solution by polarizable continuum models. J Chem Phys 109:6246–6254
144. Jalkanen KJ, Degtyarenko IM, Nieminen RM, Cao X, Nafie LA, Zhu F, Barron LD (2008) Role of hydration in determining the structure and vibrational spectra of L-alanine and N-acetyl L-alanine N′-methylamide in aqueous solution: a combined theoretical and experimental approach. Theor Chem Acc 119:191–210
145. Mukhopadhyay P, Zuber G, Beratan DN (2008) Characterizing aqueous solution conformations of a peptide backbone using Raman optical activity computations. Biophys J 95:5574–5586
146. Mukhopadhyay P, Wipf P, Beratan DN (2009) Optical signatures of molecular dissymmetry: combining theory with experiments to address stereochemical puzzles. Acc Chem Res 42:809–819
147. Nafie LA, Oboodi MR, Freedman TB (1983) Vibrational circular dichroism in amino acids and peptides. 8. A chirality rule for the methine C*H stretching mode. J Am Chem Soc 105:7449–7450
148. Oboodi MR, Lal BB, Young DA, Freedman TB, Nafie LA (1985) Vibrational circular dichroism in amino acids and peptides. 9. Carbon-Hydrogen stretching spectra of the amino acids and selected transition metal complexes. J Am Chem Soc 107:1547–1556
149. Freedman TB, Chernovitz AC, Zuk WM, Paterlini MG, Nafie LA (1988) Vibrational circular dichroism in the methine bending modes of amino acids and dipeptides. J Am Chem Soc 110:6970–6974
150. Zuk WM, Freedman TB, Nafie LA (1989) Vibrational circular dichroism in the CH-stretching modes of L-α-amino acids as a function of pH. J Phys Chem 93:1771–1779
151. Shanmugam G, Polavarapu PL (2005) Film techniques for vibrational circular dichroism measurements. Appl Spectrosc 59:673–681
152. Buschmanna H-J, Schollmeyer E, Mutihac L (2003) The formation of amino acid and dipeptide complexes with α -cyclodextrin and cucurbit[6]uril in aqueous solutions studied by titration calorimetry. Thermochim Acta 399:203–208
153. Zhang P (2006) Chiral separations and chiroptical spectroscropic studies. Dissertation, Graduate School of Vanderbilt University, Nashville
154. Jalkanen KJ, Jürgensen VW, Claussen A, Rahim A, Jensen GM, Wade RC, Nardi F, Jung C, Degtyarenko IM, Nieminen RM, Herrmann F, Knapp-Mohammady M, Niehaus TA, Friman K, Suhai S (2006) Use of vibrational spectroscopy to study protein and DNA structure, hydration, and binding of biomolecules: a combined theoretical and experimental approach. Int J Quant Chem 106:1160–1198
155. Elstner M (2006) The SCC-DFTB method and its application to biological systems. Theor Chem Acc 116:316–325
156. Barron LD, Gargaro AR, Hecht L, Polavarapu P (1991) Experimental and ab initio theoretical vibrational Raman optical activity of alanine. Spectrochim Acta 47A:1001–1016
157. Ji Z, Santamaria R, Garzón IL (2010) Vibrational circular dichroism and IR absorption spectra of amino acids: a density functional study. J Phys Chem A 114:3591–3601
158. http://chemistry.gsu.edu/Glactone/PDB/pdb.html
159. Bellamy LJ (1980) The infrared spectra of complex molecules: advances in infrared group frequencies. Chapman & Hall, New York
160. Coates J (2000) Interpretation of infrared spectra, a practical approach. In: Meyers RA (ed) Encyclopedia of analytical chemistry. Wiley, Chichester, pp 10815–10837
161. Dobrowolski JCz, Jamróz MH, Kołos R, Rode JE, Sadlej J (2007) Theoretical prediction and the first IR-matrix observation of several L-cysteine molecule conformers. ChemPhysChem 8:1085–1094
162. Dobrowolski JCz, Jamróz MH, Kołos R, Rode JE, Sadlej J (2008) IR low-temperature matrix and ab initio study on β-alanine conformers. ChemPhysChem 9:2042–2051
163. Dobrowolski JCz, Jamróz MH, Kołos R, Rode JE, Cyrański MK, Sadlej J (2010) IR low-temperature matrix, X-ray and ab initio study on L-isoserine conformations. Phys Chem Chem Phys 12:10818–10830

164. Rode JE, Dobrowolski JCz, Sadlej J (2011) Phenylisoserine in the gas-phase and water: Ab initio studies on neutral and zwitterion conformers. J Mol Model 17:961–970
165. Tanaka T, Kodama TS, Morita HE, Ohno T (2006) Electronic and vibrational circular dichroism of aromatic amino acids by density functional theory. Chirality 18:652–661
166. Fabisiewicz Z (2012) The vibrational circular dichroism of alanine in heavy water—theory and experiment, BSc Thesis under the supervision of Pecul M, and under the care of Kaminski M, University of Warsaw, Warsaw
167. Diem M, Cotkin PJ, Kupfer JM, Tindall AP, Nafie LA (1977) Vibrational circular dichroism in amino acids and peptides. 1. Alanine. J Am Chem Soc 99:8103–8104
168. Diem M, Gotkin PJ, Kupfer JM, Nafie LA (1978) Vibrational circular dichroism in amino acids and peptides. 2. Simple alanyl peptides. J Am Chem Soc 100:5644–5650
169. Diem M, Photos E, Khouri H, Nafie LA (1979) Vibrational circular dichroism in amino acids and peptides. 3. Solution- and solid-phase spectra of alanine and serine. J Am Chem Soc 101:6829–6837
170. Freedman TB, Diem M, Polavarapu PL, Nafie LA (1982) Vibrational circular dichroism in amino acids and peptides. 6. Localized molecular orbital calculations of the Carbon-Hydrogen stretching vibrational circular dichroism in deuterated isotopomers of alanine. J Am Chem Soc 104:3343–3349
171. Diem M, Polavarapu PL, Oboodi MR, Nafie LA (1982) Vibrational circular dichroism in amino acids and peptides. 4. Vibrational analysis, assignments, and solution-phase Raman spectra of deuterated isotopomers of alanine. J Am Chem Soc 104:3329–3336
172. Lal BB, Diem M, Polavarapu PL, Oboodi MR, Freedman TB, Nafie LA (1982) Vibrational circular dichroism in amino acids and peptides. 5. Carbon-Hydrogen stretching vibrational circular dichroism and fixed partial charge calculations for deuterated isotopomers of alanine. J Am Chem Soc 104:3336–3342
173. Oboodi MR, Lal BB, Young DA, Freedman TB, Nafie LA (1985) Vibrational circular dichroism in amino acids and peptides. 9. Carbon-Hydrogen stretching spectra of the amino acids and selected transition metal complexes. J Am Chem Soc 107:1547–1556
174. Diem M (1988) Infrared vibrational circular dichroism of alanine in the mid-infrared region: isotopic effects. J Am Chem Soc 110:6967–6970
175. Ellzy MW, Jensen JO, Hameka HF, Kay JG (2003) Correlation of structure and vibrational spectra of the zwitterion L-alanine in the presence of water: an experimental and density functional analysis. Spectrochim Acta 59A:2619–2633
176. Ma S, Freedman TB, Cao X, Nafie LA (2006) Two-dimensional vibrational circular dichroism correlation spectroscopy: pH-induced spectral changes in L-alanine. J Mol Struct 799:226–238
177. Freedman TB, Nafie LA (1983) Vibrational optical activity calculations using infrared and Raman atomic polar tensors. J Chem Phys 78:27–31
178. Freedman TB, Nafie LA (1983) Erratum: vibrational optical activity calculations using infrared and Raman atomic polar tensors [J. Chem. Phys. 78, 27 (1983)] J Chem Phys 79:1104
179. Nafie LA, Yu GS, Qu X, Freedman TB (1994) Comparison of IR and Raman forms of vibrational optical activity. Faraday Discuss 99:13–34
180. Gontrani L, Mennucci B, Tomasi J (2000) Glycine and alanine: a theoretical study of solvent effects upon energetics and molecular response properties. J Mol Struct (Theochem) 500:113–127
181. Tajkhorshid E, Jalkanen KJ, Suhai S (1998) Structure and vibrational spectra of the zwitterion L-alanine in the presence of explicit water molecules: a density functional analysis. J Phys Chem B 102:5899–5913
182. Frimand K, Bohr H, Jalkanen KJ, Suhai S (2000) Structures, vibrational absorption and vibrational circular dichroism spectra of L-alanine in aqueous solution: a density functional theory and RHF study. Chem Phys 255:165–194
183. Abdali S, Jalkanen KJ, Bohr H, Suhai S, Nieminen RM (2002) The VA and VCD spectra of various isotopomers of L-alanine in aqueous solution. Chem Phys 282:219–235

184. Wang F, Polavarapu PL (2001) Temperature influence on the secondary structure of avidin and avidin-biotin complex: a vibrational circular dichroism study. J Phys Chem B 105:7857–7864
185. Noda I (2008) Recent advancement in the field of two-dimensional correlation spectroscopy. J Mol Struct 883–884:2–26
186. Ryu SR, Czarnik-Matusewicz B, Dukor RK, Nafie LA, Jung YM (2012) Analysis of the molten globule state of bovine α-lactalbumin by using vibrational circular dichroism. Vib Spectrosc 60:68–72
187. Ashton L, Barron LD, Czarnik-Matusewicz B, Hecht L, Hyde J, Blanch EW (2005) Two-dimensional correlation analysis of Raman optical activity data on the alpha-helix-to-beta-sheet transition in poly(L-lysine). Mol Phys 104:1429–1445
188. Pazderka T, Kopecký V Jr (2012) Two-dimensional correlation analysis of Raman optical activity—basic rules and data treatment. Vib Spectrosc 60:193–199
189. Sadlej J (1985) Semi-empirical methods of quantum chemistry. PWN Warsaw, Ellis-Horwood, Chichester
190. Degtyarenko IM, Jalkanen KJ, Gurtovenko AA, Nieminen RM (2007) L-alanine in a droplet of water: a density-functional molecular dynamics study. J Phys Chem B 111:4227–4234
191. Mullin JM, Gordon MS (2009) Alanine: then there was water. J Phys Chem B 113:8657–8669
192. Jalkanen KJ, Suhai S, Bohr HG (2009) Quantum molecular biological methods using density functional theory. In: Encyclopedia of applied physics
193. Hecht L, Barron LD, Hug W (1989) Vibrational Raman optical activity in backscattering. Chem Phys Lett 158:341–344
194. Barron LD, Gargaro AR, Hecht L, Polavarapu PL (1992) Vibrational Raman optical activity of alanine as a function of pH. Spectrochim Acta 48A:261–263
195. Yu G-S, Freedman TB, Nafie LA, Deng Z, Polavarapu PL (1995) Experimental measurement and ab initio calculation of Raman optical activity of L-alanine and its deuterated isotopomers. J Am Chem Soc 99:835–843
196. Nafie LA (1996) Vibrational optical activity. Appl Sppectr 50:14A–26A
197. Jalkanen KJ, Nieminen RM, Frimand K, Bohr J, Bohr H, Wade RC, Tajkhorshid E, Suhai S (2001) A comparison of aqueous solvent models used in the calculation of the Raman and ROA spectra of L-alanine. Chem Phys 265:125–151
198. Kapitán J, Baumruk V, Kopecký V Jr, Bouř P (2006) Conformational flexibility of L-alanine zwitterion determines shapes of Raman and Raman optical activity spectral bands. J Phys Chem A 110:4689–4696
199. Daněček P, Kapitán J, Baumruk V, Bednárová L, Kopecký V Jr, Bouř P (2007) Anharmonic effects in IR, Raman, and Raman optical activity spectra of alanine and proline zwitterions. J Chem Phys 126:224513–526
200. Fang Y, Wua G, Wang P (2012) The asymmetry of the differential bond polarizabilities in the Raman optically active (+)-(R)-methyloxirane and L-alanine. Chem Phys 393:140–147
201. Cappelli C, Monti S, Rizzo A (2005) Effect of the environment on vibrational infrared and circular dichroism spectra of (S)-Proline. Int J Quantum Chem 104:744–757
202. Hahn S, Lee H, Cho M (2004) Theoretical calculations of infrared absorption, vibrational circular dichroism, and two-dimensional vibrational spectra of acetylproline in liquids water and chloroform. J Chem Phys 121:1849–65
203. Lee KK, Hahn S, Oh KI, Choi JS, Joo C, Lee H, Han H, Cho M (2006) Structure of N-acetylproline amide in liquid water: experimentally measured and numerically simulated infrared and vibrational circular dichroism spectra. J Phys Chem B 110:18834–18843
204. Lüdeke S, Pfeifer M, Fischer P (2011) Quantum-cascade laser-based vibrational circular dichroism. J Am Chem Soc 133:5704–5707
205. Faist J, Capasso F, Sivco DL, Sirtori C, Hutchinson AL, Cho AY (1994) Quantum cascade laser. Science 264:553–556.
206. Kapitán J, Baumruk V, Kopecký V Jr, Pohl R, Bouř P (2006) Proline zwitterion dynamics in solution, glass, and crystalline state. J Am Chem Soc 128:13451–13462

207. Qiu S, Li G, Wang P, Zhou J, Feng Z, Li C (2011) pH-dependent chirality of L-proline studied by Raman optical activity and density functional theory calculation. J Phys Chem A 115:1340–1349
208. Pecul M, Deillon C, Thorvaldsen AJ, Ruud K (2010) The aqueous Raman optical activity spectra of 4(*R*)-hydroxyproline: theory and experiment. J Raman Spectrosc 41:1200–1210
209. Pecul M (2006) Theoretical simulation of the ROA spectra of neutral cysteine and serine. Chem Phys Lett 427:166–176
210. Sadlej J, Dobrowolski JCz, Rode JE, Jamróz MH (2007) Density functional theory study on vibrational circular dichroism as a tool for analysis of intermolecular systems: (1:1) cysteine-water complex conformations. J Phys Chem A 111:10703–10711
211. Rode JE, Dobrowolski JCz (2003) VCD technique in determining intermolecular H-bond geometry. A DFT study. J Mol Struct (Theochem) 637:81–89
212. Dobrowolski JCz, Rode JE, Sadlej J (2012) VCD chirality transfer: a new insight into the intermolecular interactions. In: Leszczynski J, Shukla MK (eds) Practical aspects of computational chemistry. Springer, Netherlands, pp 451–478
213. Zhu P, Yang G, Reza Poopari M, Bie Z, Xu Y (2012) Conformations of serine in aqueous solutions as revealed by vibrational circular dichroism. ChemPhysChem 13:1272–1281
214. Deplazes E, van Bronswijk W, Zhu F, Barron LD, Ma S, Nafie LA, Jalkanen KJ (2008) A combined theoretical and experimental study of the structure and vibrational absorption, vibrational circular dichroism, Raman and Raman optical activity spectra of the L-histidine zwitterion. Theor Chem Acc 119:155–176
215. Zhu P (2012) Conformations of some amino acids in aqueous solutions by vibrational circular dichroism spectroscopy. MSc Dissertation under supervision of Xu Y, University of Alberta, Edmonton
216. Reza Poopari M, Zhu P, Dezhahang Z, Xu Y (2012) Vibrational absorption and vibrational circular dichroism spectra of leucine in water under different pH conditions: hydrogen-bonding interactions with water. J Chem Phys 137:194308–194314
217. Rai AK, Xu X, Lin Z, Rai DK (2011) Conformational search for zwitterionic leucine and hydrated conformers of both the canonical and zwitterionic leucine using the DFT-CPCM model. Vib Spec 56:74–81
218. Rode JE, Dobrowolski JCz, Sadlej J Prediction of L-Methionine VCD Spectra in the Gas Phase and Water Solution. J. Phys. Chem. B, DOI: 10.1021/jp405462a
219. Chruszcz-Lipska K, Czernicki W, Kaczor A, Barto Z (2012) Raman optical activity spectrum of L-valine: experiment and theory. In Abstracts of 23rd International Conference on Raman Spectroscopy, Bangalore—ICORS 2012, 12–17 August 2012
220. Jacob CR, Luber S, Reiher M (2008) Calculated Raman optical activity signatures of tryptophan side chains. ChemPhysChem 9:2177–2180
221. Blanch EW, Hecht L, Day LA, Pederson DM, Barron LD (2001). Tryptophan absolute stereochemistry in viral coat proteins from Raman optical activity. J Am Chem Soc 123:4863–4864
222. Luber S, Reiher M (2009) Intensity-carrying modes in Raman and Raman optical activity spectroscopy. ChemPhysChem 10:2049–2057
223. Ruud K, Taylor PR, Åstrand PO (2001) Zero-point vibrational effects on optical rotation. Chem Phys Lett 337:217–223
224. Bonmarin M (2010) Development of ultrafast time-resolved chiral infrared spectroscopy, PhD Dissertation, Mathematisch-naturwissenschaftlichen Fakultät, der Universität Zürich, Züurich
225. Bonmarin M, Helbing J (2008) A picosecond time-resolved vibrational circular dichroism spectrometer. Opt Lett 33:2086–208
226. Helbing J, Bonmarin M (2009) Time-resolved chiral vibrational spectroscopy. Chimia 63:128–133
227. Bonmarin M, Helbing J (2009) Polarisation control of ultrashort mid-IR laser pulses for transient vibrational circular dichroism measurements. Chirality 21:298–306
228. Helbing J, Bonmarin M (2009) Vibrational circular dichroism signal enhancement by self-heterodyning using elliptically polarised laser pulses. J Chem Phys 131:174507–174516

229. Polavarapu PL, Shanmugam G (2011) Comparison of mid-infrared fourier transform vibrational circular dichroism measurements with single and dual polarization modulations. Chirality 23:801–807
230. Lakhani A, Malon P, Keiderling TA (2009) Comparison of vibrational circular dichroism instruments: development of a new dispersive VCD. Appl Spectr 63:775–785
231. Nafie LA (1996) Theory of resonance Raman optical activity—the single electronic-state limit. Chem Phys 205:309–322
232. Vargek M, Freedman TB, Lee E, Nafie LA (1998) Experimental observation of resonance Raman optical activity. Chem Phys Lett 287:359–364
233. Luber S, Neugebauer J, Reiher M (2010) Enhancement and de-enhancement effects in vibrational resonance Raman optical activity. J Chem Phys 132:044113–044129
234. Johannessen C, White PC, Abdali S (2007) Resonance Raman optical activity and surface enhanced resonance Raman optical activity analysis of cytochrome C. J Phys Chem A 111:7771–7776
235. Yamamoto S, Bouř P (2012) Detection of molecular chirality by induced resonance Raman optical activity in Europium complexes. Angew Chem Int Ed Engl 51:11058–11061
236. Jensen L, Autschbach J, Krykunov M, Schatz GC (2007) Resonance vibrational Raman optical activity: a time dependent density functional theory approach. J Chem Phys 127:134101–134112
237. Nafie LA (2010) Resonance Raman optical activity: past, present and future. Proceedings of XXII International Conference on Raman Spectroscopy 1267:98–98, Boston, USA, 8–10 August 2010
238. Merten C, Li H, Nafie LA (2012) Simultaneous resonance Raman optical activity involving two electronic states. J Phys Chem A 116:7329–7336
239. Kapitán J, Barron LD, Hecht L (2012) Raman optical activity spectrometer in the deep UV spectral region. In: Abstracts of 23rd International Conference on Raman Spectroscopy, Bangalore, August 12–17, 2012
240. Lal S, Alexander M (2010) MURI Final Report: a multimodality ultramicrospectroscope (MUMS) nanoscale imaging with integrated spectroscopies for chemical and biomolecular identification, William Marsh Rice University, Huston
241. Unno M, Shingae T, Kikukawa T, Kamo N (2010) Exploring the active site structure of photoreceptor proteins by Raman optical activity. Proceedings of XXII International Conference on Raman Spectroscopy 1267:102–103, Boston, 8–10 August 2010
242. Abdali S, Blanch EW (2008) Surface enhanced Raman optical activity (SEROA). Chem Soc Rev 37:980–992
243. Osińska K, Pecul M, Kudelski A (2010) Circularly polarized component in surface-enhanced Raman spectra. Chem Phys Lett 496:86–90
244. Janesko BG, Scuseria GE (2006) Surface enhanced Raman optical activity of molecules on orientationally averaged substrates: theory of electromagnetic effects. J Chem Phys 125:124704–124716
245. Jensen L (2009) Surface-enhanced vibrational Raman optical activity: a time-dependent density functional theory approach. J Phys Chem A 113:4437–4444
246. Novák V, Šebestík J, Bouř P (2012) Theoretical modelling of the surface-enhanced Raman optical activity. J Chem Theory Comput 8:1714–1720
247. Brazhe NA, Brazhe AR, Sosnovtseva OV, Abdali S (2009) Novel chiroptical analysis of hemoglobin by surface enhanced resonance Raman optical activity spectroscopy. Chirality 21:E307–312
248. Johannessen C, Abdali S (2007) Surface enhanced Raman optical activity as an ultra sensitive tool for ligand binding analysis. Spectroscopy 21:143–149
249. Pour SO, Bell SEJ, Blanch EW (2011) Use of a hydrogel polymer for reproducible surface enhanced Raman optical activity (SEROA) Chem Commun 47:4754–4756
250. Li H, Nafie LA (2012) Simultaneous acquisition of all four forms of circular polarization Raman optical activity: results for α-pinene and lysozyme. J. Raman Spectrosc 43:89–94

Chapter 6
Electronic Circular Dichroism Spectroscopy in Structural Analysis of Biomolecular Systems

Magdalena Pecul and Wojciech Dzwolak

Abstract In this chapter, the use of electronic circular dichroism (CD) in structural analysis of biological molecular systems is discussed. The quantum chemical theory of the phenomenon is briefly presented, and the applications of the CD spectroscopy in determination of spatial structure of proteins and nucleic acids, and of absolute configuration of small molecules of pharmacological significance are summarized. Some attention is also paid to the emerging technique related to CD—circularly polarized luminscence (CPL).

(Dated: 1 February 2013)

Keywords Electronic circular dichroism • Circularly polarized luminscence • Protein structure • Nucleic acid structure • Induced chirality

6.1 Introduction

Since the majority of building blocks of living matter are chiral, chiroptical methods have a special significance in structural analysis of compounds of biochemical relevance. Electronic circular dichroism (ECD, usually abbreviated in the literature to CD, after circular dichroism) is the most widespread among them, although vibrational chiroptical spectroscopic methods: vibrational circular dichroism and Raman optical activity are being developed very rapidly nowadays.

Dependence of the optical rotation on the incident light wavelength was discovered already in 1817 by Biot [1], but this knowledge was not utilized until mid-XX century, when optical rotatory dispersion (ORD) measurements started to be routinely conducted [2, 3]. Circular dichroism spectroscopy has a similar history: the phenomenon was first observed in 1847 [4] and the connection between ORD and CD was realized by Cotton half a century later [5], but it became a widespread spectroscopic technique only about 1960 [3]. (Early days of ORD and CD have been reviewed recently by Laur [6].) Since the 1960s, CD has not only replaced optical rotatory dispersion (ORD) as the main biophysical tool to probe static conformations of biopolymers in solution, but it has also successfully evolved into an

M. Pecul (✉) · W. Dzwolak
Faculty of Chemistry, University of Warsaw, Pasteura 1,
02-093, Warszawa, Poland
e-mail: mpecul@chem.uw.edu.pl

M. Baranska (ed.), *Optical Spectroscopy and Computational Methods in Biology and Medicine,* Challenges and Advances in Computational Chemistry and Physics 14,
DOI 10.1007/978-94-007-7832-0_6, © Springer Science+Business Media Dordrecht 2014

insightful experimental approach in time-resolved studies of protein folding, as well as in the field of protein-ligand interactions.

There are plenty of excellent reviews on applications of CD in protein chemistry and biophysics—either as book chapters [7–15] or up-to-date review papers [16–20]. Out of necessity, in this concise work, we will focus on some general aspects of applications of CD in broadly understood bioscience. First, we discuss briefly the quantum chemical theory underlying the spectroscopic phenomenon and the main computational methods used for modeling the spectra. Then, we proceed to outlining the use of CD in obtaining structural information on peptides, proteins and nucleic acids. The main emphasis is on new developements in this area, and in particular on the use of Synchrotron Radiation Circular Dichroism on the one hand, and on induced CD of achiral chromophores (in near UV or visible range) on the other. Although the main applications of CD in bioscience are in the studies of biopolymers, we have decided to devote some attention also to the use of CD in establishing absolute configuration of small chiral molecules, since many of these studies concern drugs or potential drugs, and therefore are relevant to the topic. Finally, we discuss an emerging technique related to CD—circularly polarized luminscence (CPL)—which enables probing the structure of excited states of chiral molecules.

6.2 Theory and Modelling of CD

Along with extending the range of biochemical applications of electronic CD spectra, understanding of the relationship between structural properties and CD spectrum has also been deepened by relatively recent developements in theoretical modelling of CD spectra.

The quantum chemical theory of the CD phenomenon has been described as early as in 1928 by L. Rosenfeld [21], who related the difference between the molar extinction coefficients for left and right circularly-polarized light, e_L(A)—e_R(A), to the rotatory strength ^{n}R of the transition between the ground state 0 and the nth excited state. The rotatory strength ^{n}R was shown for isotropic samples to be given as the product of the electric dipole transition moment (0|$\boldsymbol{\mu}$|n) and magnetic dipole transition moment (n|$\mathbf{m}$|0). In atomic units and Dirac notation, it can be written as:

$$^{n}R = \langle 0 \,|\, \boldsymbol{\mu} \,|\, n \rangle \cdot \langle n \,|\, \mathbf{m} \,|\, 0 \rangle \tag{6.1}$$

For oriented samples, there is also an analogous contribution from interactions with the electronic quadrupole moment, containing a product of electric dipole transition moment and electric quadrupole transition moment, but it is purely anisotropic and therefore vanishes upon orientational averaging in isotropic samples without contributing to the measured CD. However, in principle it should be accounted for in the case of partially oriented samples (a not infrequent situation for chromophors in biomolecules, for example large fibrous proteins).

Calculations of ECD spectra for sizeable molecules of biological significance by means of ab initio quantum chemical methods have become possible due to

the developement of time-dependent density functional theory [22–27] (the time-dependent Hartree-Fock approach, developed earlier, is generally not suitable for handling excited states and the wave function methods including electron correlation are too costly in terms of computational resources), and soon afterwards they have been carried out for molecules of biological significance, including amino acids [28, 29], and for small peptides [30–32].

Nowadays, the majority of ab initio quantum chemical calculations of the CD spectra are carried out by means of DFT. There are two main factors influencing the quality of the results of such calculations: a selection of a basis set and of an exchange-correlation functional. While there is a general consensus that the basis set suitable for calculations of the CD spectra should be similarly built as for other optical properties (a balanced valenced set and diffuse functions are necessary), the choice of an exchange-correlation functional is more problematic, especially for transitions to diffuse Rydberg states. Some authors advocate the use of range-separated functionals for the purpose [33], but the issue is still debated. Assuring gauge invariance, a major problem in calculations of the optical rotation, is a not a critical issue in the case of CD, since the results obtained using the velocity gauge and the length gauge with and without gauge-including orbital are very similar [34].

A lot of effort is devoted to extending the range of systems tractable by means of ab initio methods (for example through developement of linear scaling methods, and adaptation of GPU computing for the purpose). Still, most of the chiral molecules of biochemical and biological interest are far too large to be handled by means of ab initio quantum chemical calculations. Thus, at the present, the only feasible method to utilize quantum chemistry in calculations of ECD spectra for sizable peptides is to carry out the calculations for one chromophore unit and to simulate the final spectrum by treating the biopolymer as an assemble of chromophores and transfering the optical properties calculated for individual chromophores to this assemble [35, 36]. This can be done by means of several different methods. The most approximate is the dipole interaction model [37], which considers individual atoms and the amide chromophore to be point dipole oscillators. In the presence of an electric field, they interact through mutually induced dipole moments. Another approach to calculating protein CD is the matrix method [38, 39], where a Hamiltonian matrix is constructed assuming the protein consists of non-interacting chromophoric groups. This matrix is then diagonalized by a unitary transformation, which describes the transition from a non-interacting regime to an interacting regime. Using this transformation, the properties calculated for individual chromophores (for example amide groups) are transformed and the CD of the protein is computed. In some cases, meaningful information can even be obtained from a simplified exciton coupling approach [40–42], in which the rotatory strength of the coupled system is calculated from the electric dipole transition moments and the position vector between them.

Most of the simulations of the CD spectra of proteins are focused on modelling of interacting amide chromphores. An isolated amide chromphore is usually approximated by N-methylacetamide (see for example Refs. 43–45), and a number of succesful simulations of the CD spectra of peptides have been carried out by

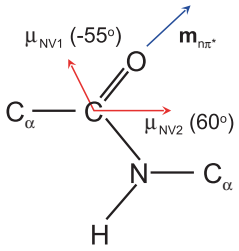


Fig 6.1 Directions of the nπ* magnetic dipole transition moment (blue arrow) and ππ* electric dipole transition moments (red arrows) in the peptide group, according to Sreerama and Woody [7]

means of this model [20, 35, 36, 38, 46]. Other chromophores contributing to the CD spectrum of peptides and proteins are aromatic groups. For them, the transition dipoles are nearly additive, and they are very suitable for calculations using the exciton coupling model. Since they can provide unique structural information (see below), they have been the subject of a number of computational studies [20, 47, 48]. The CD spectrum of the disulphide chromophore, although less important in practice, has also been modeled [29, 49, 50].

6.3 Applications of CD in Proteomics

The ubiquitous left-handedness of all except one (glycine) translational amino acids, when coupled to the electronic transitions in amide chromofore in the peptide bond (Fig. 6.1), gives rise to characteristic CD bands whose shape and magnitude are intimately linked to the conformation of the polypeptide chain. While the corresponding spectra in the so-called far-UV range (ca. 190–250 nm) are most informative in terms of protein secondary structure, and also represent the most conventional biophysical application of CD, changing the spectral range may highlight different electronic transitions probing distinct aspects of protein structure and its dynamics. Table 6.1 provides a brief summary of the spectral transitions that can be captured at different wavelengths in protein CD spectra and which originate from different parts of protein structure. We will review the applications of CD in proteomics starting from the signals of the amide chromophores (in far UV), then discussing aromatic amino acids chromophores (in near UV) and finally the induced CD from achiral chromophores (in near UV and visible light).

A. CD signals of amide chromophores Observation of high-energy electronic transitions in the vacuum-UV range often calls for synchrotron radiation sources (thus the term Synchrotron Radiation Circular Dichroism—SRCD—was coined) and due to the necessity of overcoming a number of apparatus- and sample-handling-related problems, is still far from being routine. Initially, the low-wavelength bands detected in protein vacuum-UV CD spectra were incorrectly assigned to amide nΣ* transitions [51]. Although their origins are being still debated, contributions from charge-transfer transitions from (i) oxygen lone pair to peptide π* orbital, or (ii) from nonbonding n orbital on one peptide group to π* orbital on the

Table 6.1 Wavelength range of electronic CD spectra of proteins

Types of CD	Wavelength range	Electronic transition	Application	Ref.
Vacuum-UV CD/SRCD	140–190 nm	Charge-transfer nπ*/ππ*	Improved estimates of secondary structure compared to far-UV only CD	19, 52–54, and 104
Far-UV CD	190–250 nm	Peptide group nπ*/ππ*	Conventional analysis of secondary structure	7–9, 16–18
Near-UV CD	250–300 nm	Aromatic side chains nπ*/ππ*, disulfide nΣ*	Detection of tertiary structure	7 and 105
Visible CD (Extrinsic Cotton effect)	350–700 nm	Excitation of prosthetic groups	Probing ligand-protein interactions	18, 106, and 107

neighboring group are likely [52]. Most importantly, taking into consideration the corresponding vacuum-UV CD bands along with conventional far-UV CD bands results in a significant improvement of accuracy of CD-based estimation of protein secondary structure components [53]. It is also argued that expanding the far-UV of protein CD spectra with shorter wavelength range may enable distinguishing between certain secondary structures such as parallel vs. antiparallel β-sheets, or α- vs. 3_{10}-helices which is otherwise impossible through the conventional CD [53].

The most commonly studied far-UV CD spectral range is dominated by nπ* and ππ* transitions within the peptide group, although these can be overlapped to some extent by signals from aromatic side-chains in proteins with a high abundance of tryptophan, tyrosine, and phenylalanine (e.g. Ref. 17). On the other hand, for the low-intensity far-UV signal of polyproline II conformation the contamination from the high-energy transitions in vacuum-UV may be significant [54]. Still, for a typical protein, the corresponding far-UV CD spectrum will reflect sum of contributions from different secondary components, and these could be approximated through simplified spatial arrangements of main chain amide chromophores. Far-UV CD spectra of various types of secondary structure are shown in Fig. 6.2. The characteristic spectral features of α-helix—the most common secondary motif in proteins—consist of a couplet of strong positive and negative signals at 192 and 208 nm, respectively (both stemming from ππ* transitions), and another negative and partly overlapping peak at 222 nm which originates from nπ* transitions. CD intensity of a helical stretch of a polypeptide chain is approximately proportional to its length (i.e. the average ellipticity of a helix per amino acid residue is constant). However, with decreasing length of a helix, a deviation from this rule caused by frayed unstructured ends of the helix becomes apparent. Simple empirical equations have proven adequate to correct for this effect [7].

In contrast to the sharply defined geometrical constrains of an α-helix, β-sheets are much more diverse equally in terms of arrangement of strands (parallel vs. antiparallel), their local twist, and a strength of interstrand hydrogen bonds. The adaptability of the β- conformation to changes in local geometry is reflected by the large

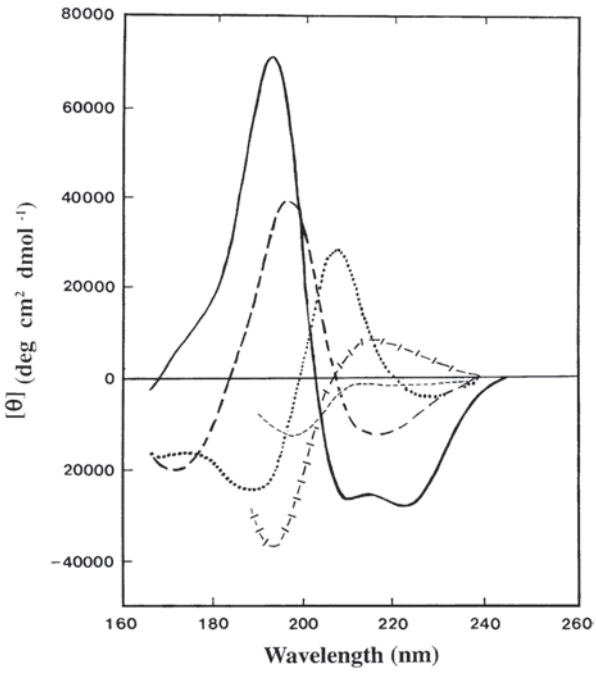


Fig. 6.2 Far UV CD spectra associated with various types of secondary structure. Solid line, α- helix; long dashed line, anti-parallel (β-sheet; dotted line, type I (β-turn; cross dashed line, extended 31-helix or poly (Pro) II helix; short dashed line, irregular structure, according to Kelly, Jess and Price [18]

area on the Ramachandran plot assigned to β-sheets. This structural polymorphism of extended structures, especially those aggregated in the form of so-called amyloid fibrils, may lead to pronounced differences in corresponding far-UV CD spectra [55, 56]. On the other hand, the CD signal of a typical β-sheet structure found in globular proteins yields a rather simple spectrum containing a positive band at ca. 195 nm (ππ*) and a negative one at 216 nm (nπ*), as shown in Fig. 6.2.

The polyproline II conformation is thought to represent a major share of unfolded or disordered protein state. It is also assumed by charged homopolypeptide chains such as poly-L-lysine and poly-L-glutamic acid at neutral pH. The corresponding CD spectra are rather flat with a stronger negative band at 200 nm and a very weak positive one near 218 nm. Typically, a gradual denaturation of α-helical protein e.g. by temperature will be followed in far-UV CD spectra by a stepwise decay of the 208 nm/222 nm negative bands and a simultaneous appearance of the spectral features of the emerging polyproline II conformation. Certainly, denaturation or even covalent degradation of a protein does not have to result in a complete breakdown of the native structure. Still, far-UV CD spectroscopy has proven sensitive to follow such limited or localized changes in the native state, as has been recently shown in a time-lapse spectroscopic study on enzymatically-degraded β-lactoglobulin and β-casein [57].

Finally, the variety of turns present in protein molecules will also contribute to the net far-UV CD spectrum. Because soluble polypeptide model with a neat turn structure of a particular type are often scarce or inaccessible, assignment of

different turn geometries to CD bands often lacks solid empirical verification. The problem has been briefly addressed in earlier works (e.g. Ref. 7) and is investigated by means of quantum chemical calculations [58].

Because of the near additivity of these signals, a far-UV CD spectrum of a protein with a known structure can be relatively easily predicted (as discussed above), given the typical structural similarity of crystal and water-dissolved proteins. From a practical standpoint, a much more interesting and important is the opposite goal: estimation of the content of various secondary structure elements based on an experimentally collected far-UV CD spectrum. Achieving it permits using CD spectroscopy as a very fast and quantitative means of analyzing protein secondary structure in solution. The study by Greenfield and Fasman [59] is one of the earliest and most straightforward approaches to this problem. The authors used far-UV CD spectra of poly-L-lysine—a homopolypeptide capable of adopting (depending on pH and temperature) either α-helical, antiparallel β-sheet, or random coil (polyproline II) conformation—as a reference set for semi-quantitative deconvolution of several globular proteins. Since that seminal work it has become clear that a more accurate far-UV CD- based estimation of ingredient secondary components would require both more advanced computational algorithms and more expanded reference set of CD spectra of proteins with well-defined (through X-ray diffraction or solution NMR) 3D-structure. Several existing tools (including internet-accessible online servers [60]) for quantitative analysis of the far-UV CD spectra of proteins have been described and reviewed earlier [61–63]. As was mentioned in the previous paragraphs, expanding the CD spectral range to include vacuum-UV-detectable transitions markedly improves accuracy of the method, however, this also requires having a protein CD reference set properly expanded in the range of low wavelengths [53, 64–66].

B. CD signals from aromatic rings In the absence of electronic transitions stemming from prosthetic groups (such as heme) CD signals of proteins in the near-UV region (which could be arbitrarily defined as 250–300 nm) originate from disulfide bridges and aromatic side chains (Table 6.1). Typically, these are weaker than far-UV CD by at least an order of magnitude. Diagnostic value of broad and featureless disulfide bands is negligible. On the other hand, CD of the aromatic bands did find specific application in biophysics. The coupling of achiral chromophores of phenylalanine, tyrosine, or tryptophan planar side-chains to nearest chiral carbon atom is weak and the corresponding low rotational strength in a fluctuating disordered protein chain would be sufficiently low to make the CD signal almost undetectable. The situation changes when upon 'hydrophobic effect'-driven folding the non-polar amino acid side chains (and the aromatic residues among them) are clustered and buried in a protein hydrophobic core. Transition dipole moments of the individual aromatic chromophores interact strongly, forming a chiral quasi-united chromophore with a rather strong CD signal [7, 67]. Hence, the tertiary structure of a stable unambiguously defined protein envelope surrounding the aromatic residues is a condition sine qua non of the occurrence of measurable near-UV CD. Such CD signal would be only yet-another (next to the far-UV CD) chiroptical aspect

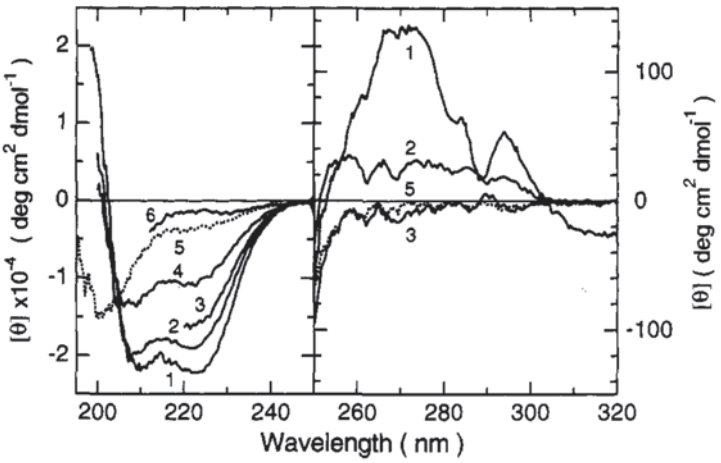


Fig. 6.3 Far-UV (left panel) and corresponding near-UV (right panel) CD spectra of various conformational states for myoglobin and apomyoglobin at 20 °C: (1) native-state myoglobin at pH 6.0; (2) native-state apomyoglobin at pH 6.0; (3) molten globule state apomyoglobin stabilized by 20 mM sodium trichloroacetate at pH 2.0 (10 mM HCl); (4) molten globule state apomyoglobin stabilized by0.4 M sodium trichloroacetate at pH 2.0 (10 mM HCI); (5) acid-unfolded state apomyoglobin at pH 2.0 (10 mM HCl); (6) unfolded state apomyoglobin in 4.0 M Gdn-HC1 at pH 6.0, according to Nishii et al [108]

of the native conformation if the native-disordered transitions were always taking place cooperatively at every level of hierarchical architecture of protein native state. However, detailed analysis of gradual denaturation of certain proteins, such as α-lactalbumin, carried out simultaneously in the far-UV and near-UV CD spectral ranges revealed that the unfolding may be non-cooperative i.e. an intermediate state (so-called molten globule state [68]) lacking native tertiary structure but with native-like secondary structure may be formed. This is illustrated by the CD data on acidic denaturation of apomyoglobin in Fig. 6.3.

The main benefit of protein CD investigations in the near-UV range is that they provide an independent insight into tertiary rather than secondary-level stability of protein. Time-resolved or time-lapse studies research focused on protein folding/unfolding probed in both the far- and near-UV CD regions are the most reliable biophysical tools to detect molten globule type intermediate states. It must be stressed, however, that very little may be ascertained in terms of type of the tertiary structure on the basis of only the shape and intensity of a single near-UV CD spectra. Even very minor variation in composition and 3D-packing of the aromatic residues in the hydrophobic core may cause dramatic changes in the sign, shape, and intensity of the CD signal: for instance, in the near-UV CD spectra of lactalbumin and lysozyme—two closely related and structurally homological α/β proteins—have opposite signs [69, 70].

C. Induced CD of achiral chromophores CD signal of ordered aromatic residues could be conceptualized as induced CD, i.e. a Cotton effect that is not an inherent

property of an asymmetric molecular structure wherein the chromofore and the chiral center are covalently fixed and coupled, but rather a consequence of complex intramolecular transition which favors specific chiral rotamers with strong CD. A similar situation may take place when an achiral chromophore binds to protein and the docking interactions enforce a chiral and therefore optically-active conformation in the the ligand. This induced circular dichroism (ICD, also termed extrinsic Cotton effect) is usually observed within near-UV or even visible range of light i.e. at the ligands specific electronic transition (Table 6.1). Certainly, induced CD can be also triggered under more subtle circumstances—for instance when upon the binding event the ligand remains rigid and achiral in terms of its covalent structure but enters chiral fields of the protein, etc. [71]. Examples of applications of ICD spectroscopy for detection and characterization of protein-ligand interactions in solution include binding of bilirubin [72] and retinoic acid [73] to β-lactoglobulin. Another example is ICD of thioflavin T (ThT). ThT is an important ligand for amyloid fibrils—linear aggregates of misfolded protein molecules which form in vivo in the course of certain neurodegenerative maladies, such as Alzheimer disease. The dye has become the most widely used amyloid-specific fluorescent stain, which according to the prevailing model of amyloid-ThT interactions binds specifically to multistranded β-sheets, such as those constituting amyloid fibrils. In solution, ThT rotates freely not only as a whole, but also internally around its single C-C bond linking benzothiazole and dimethylaminobenzene rings. In consequence, time-averaged conformation of a ThT molecule has effectively the Cs symmetry and is therefore achiral. We have shown earlier that a free ThT molecule has two mirror-image equilibrium structures at the dihedral angle of ±37°. Stabilization of one of them by of a chiral environment induces optical activity in the dye [74], as all ThT conformers are chiral except for those perfectly planar, or twisted at the right angle. This has triggered the interest in using ICD of amyloid-bound ThT as a spectroscopic probe of chiral bias of amyloid surface moieties [75–78]. A very interesting property of ThT emerged when the same ICD methodology was used to study interactions of the stain with α-helical poly-L-glutamic acid—an amyloidogenic peptide in a non-fibrillar conformation, which according to the widespread notion was not supposed to bind the amyloid-specific dye. The data shown in Fig. 6.4 shows, however, that ThT does bind to this non-amyloid conformation, as is revealed by ICD spectra. It should be stressed that the binding could not be observed through the fluorescence emission typically employed to detect ThT-protein binding, as the interactions with α-helical poly-L-glutamic acid do not induce enhancement of the quantum yield of ThT fluorescence.

In this brief subsection we have mentioned a number of applications of electronic CD in respect to the spectral wavelength range employed. CD brings information not only on the main chain structure, but also on more subtle effects of packing intrinsic aromatic chromophores of protein (ruled by the presence of tertiary structure) and extrinsic chromophores binding through non-covalent interactions. We will show that CD studies on DNA could be categorized in a similar order.

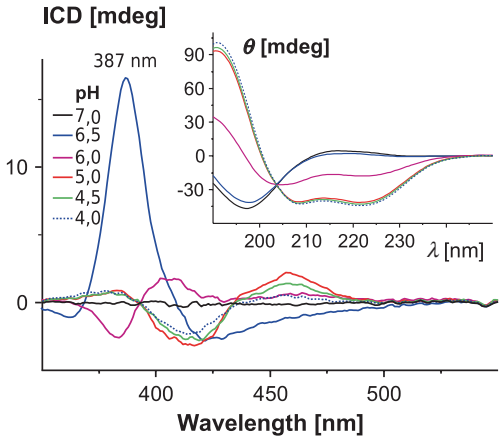


Fig. 6.4 pH-Dependence of ICD spectra of ThT in the presence of PLGA at 25 °C. Inset shows corresponding far-UV CD spectra reflecting the simultaneous disordered-to-α-helix transition, according to Babenko and Dzwolak [109]

6.4 CD of Nucleic Acids

The local molecular chirality of nucleic acids originates from D-sugar molecules (ribose in RNA and deoxyribose in DNA) which produce CD bands in remote vacuum-UV. Although purine and pyrimidine bases have many electronic transitions (mainly ππ* overlapping weaker nπ*) in an easily accessible UV range, their perfect planarity calls for additional conditions to be fulfilled before the corresponding spectral bands could become CD-active. Weak coupling to the sugar chiral center allows only for a small CD effect. However, the helical structure of DNA enables building a strong and conformation-sensitive Cotton effect. The double-helix structure of DNA maintains large numbers of closely interacting bases arranged in a chiral staircase. In an analogy to the case of induced CD in near-UV CD emerging from chiral arrangement of achiral aromatic amino acid side chains in proteins, CD bands of helical DNA (or RNA) lying between 200 and 300 nm follow from asymmetry of the overall structure of the folded biopolymer. In case of far-UV protein CD, the spectral signal reflects primarily the conformation, and the amino acid sequence affects it only indirectly (given that partly overlapping but scarce aromatic side chains cause only negligible contribution to UV absorption below 250 nm) by promoting a particular secondary structure. Basically, the amide chromophore is the same regardless of the sequence (one exception being proline). This is in contrast to the more intricate case of DNA CD where the contributing chromophore (bases) and corresponding electronic transitions contributing to CD change with the sequence of the nucleic acid. Hence fundamentally similar conformations of DNA samples with different abundances of GC and AT pairs will have distinct CD spectra. Because local interactions of transition dipoles between neighboring bases affect chiroptical characteristics of folded nucleic acid, the sequence rather than just nucleotide composition has to be taken into account when one attempts to simulate the CD spectrum of DNA with a known structure—a task remarkably more difficult than computing protein far-UV CD.

The subject of applications of CD in DNA research has been treated in detail in a number of excellent review works [9, 79–82], and our current paper is only a brief attempt to highlight certain similarities and differences in regard to protein CD. Just like in the case of proteins, locally and globally distinct conformations may be accessible to a DNA chain—depending on its length, nucleotide sequence, and physicochemical conditions. DNA melting is a phenomenon of temperature-induced unfolding of the double-helix structure and an obvious equivalent of thermal denaturation of a protein. Because the melting causes disarray of the helical staircase holding nucleotide chromophores it is reflected by a simultaneous decrease of the Cotton effect. However, temperature-induced denaturation is only an iceberg tip of plethora of fascinating applications of CD in nucleic acid research, most of which are related to polymorphism of DNA helices. The right-handed B-type double helix is the most common form of DNA and one out of several structures that can be induced in vitro by varying co-solvent concentration and ionic strength in DNA solutions. Transitions between these different helical structures may be conveniently probed by CD (Fig. 6.5). The naturally-occurring B-form converts into likewise right-handed but of larger diameter A-form in the presence of certain organic solvents, such as trifluoroethanol (TFE). In CD spectra of CG- rich DNA this is typically reflected by a shift of main positive band from ca 275 nm to 260 nm and simultaneous emergence of a negative band below 220 nm. Unlike A- and B-type double helices, the Z-form of DNA is a left-handed helix. It may be formed from the B-type in presence of high salt concentrations or cosolvents (depending on sequence). The reversal of helical sense must cause pronounced changes in the CD spectra, but their character again depends on the exact covalent nature of the DNA chain (e.g. Ref. 82 vs. 83). In fact, the A, B, and Z forms of double helix are the simplest aspects of DNA polymorphism. Circular dichroism has been successfully utilized to probe other nonclassical structures of DNA such as triplex, quadruplex, or parallel DNA. Expert review works (for example that by Maurizot [80]) treat these problems in detail.

Perturbations in DNA conformation may follow from more physiological factors than drastic changes in solvation that are often used in vitro—for example through docking interactions with proteins or small ligands specifically binding to nucleic acids. There are excellent reviews devoted to application of CD to protein-DNA interactions available [13, 81] and therefore this complex subject will not be discussed in this brief review.

6.5 CD and Absolute Configuration of Biologically Active Molecules

For many years, CD has been used to establish unknown absolute configurations of synthetic or naturally occuring chiral compounds (in many cases with some biological significance, for example pharmaceutical potential). At first, such assignment relied on empirical rules such as quadrant or octant rule, which naturally limited the range of application of CD in this field. With advancement of theoretical

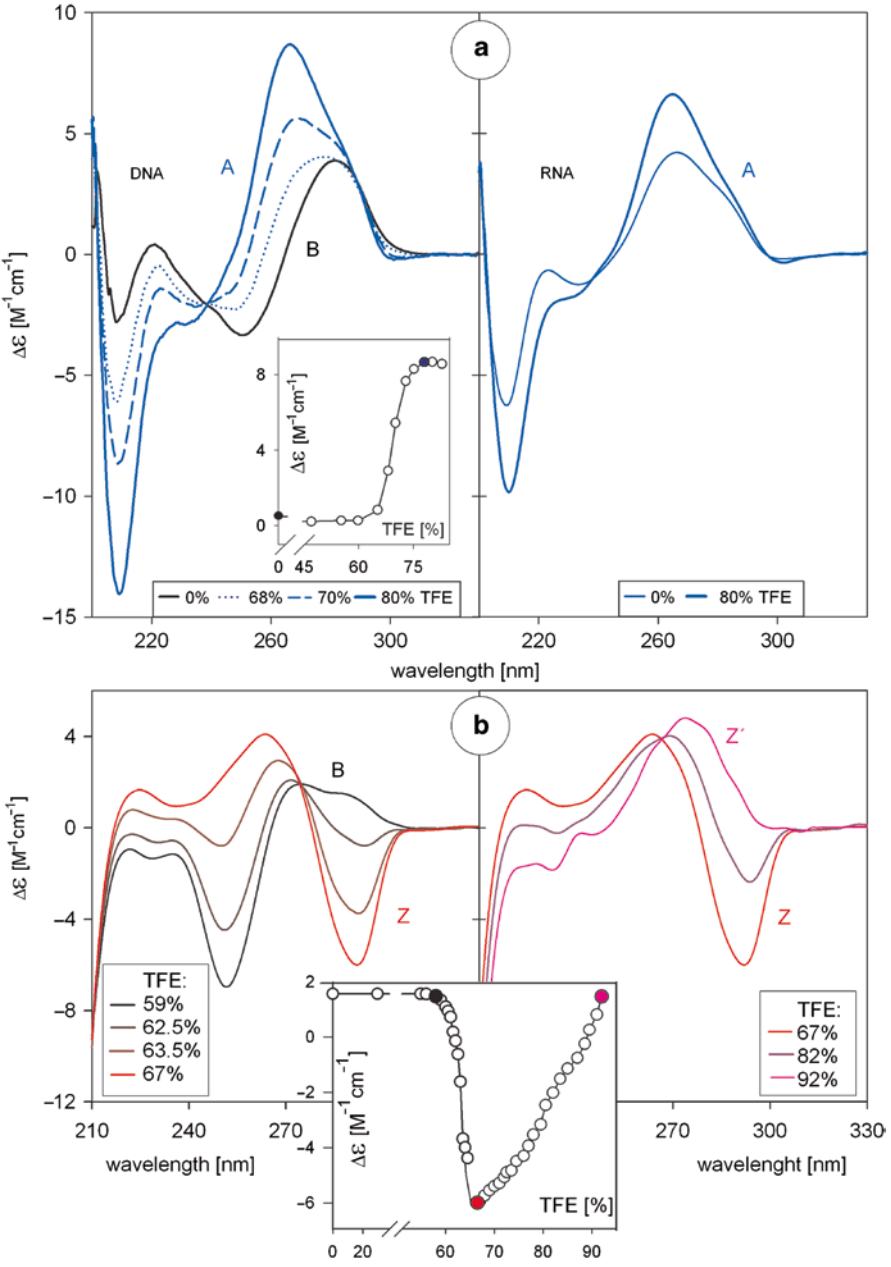


Fig. 6.5 BA and BZ transitions of DNA. (A) Left panel: CD spectra reflecting trifluorethanol-induced BA transition of the d(GCGGCGACTGGTGAGTACGC) duplex. Insert: the transition monitored at 266 nm. Right panel: CD spectra of RNA of the same sequence (U instead of T) duplexed with complementary DNA strand. (B) CD spectra reflecting trifluorethanol-induced BZ (left panel) and ZZ0 (right panel) transitions of poly[d(GC)] duplex. Insert: the transitions monitored at 291 nm. In this figure, TFE was added to the oligonucleotides dissolved in 1mM Na phosphate, 0.3 mM EDTA, pH 7 and CD spectra were measured at 0 °C [82]

calculations, a protocol has been established where the spectrum is first calculated (by means of semi-empirical or ab initio methods) for a given configuration (or a number of configurations, if several chiral centres exist) and then, band by band, compared with an experimental spectrum. An agreement of the signs between the calculated and measured spectrum is treated as an indication that the configuration has been assigned correctly.

This method is very useful, especially in cases when a crystal suitable for X-ray measurements cannot be obtained, but theoretical prediction of a CD spectrum may not be accurate for a number of reasons, which would affect the assignment. First, all low-energy conformations should be accounted for with correct weights, since the CD spectrum is very sensitive to molecular conformation. Second, some of the bands (especially those corresponding to Rydberg transitions) may change signs in different solvents, an effect which is difficult to reproduce in silico. Third, a calculation of an electronic CD even for an isolated rigid molecule is a challenging task in itself, especially when the size of a molecule precludes the use of many-configurational methods, and the only ab initio method which can be used is non-hierarchical DFT. An interesting example of the use of electronic CD for determination of the absolute configuration of a series of compounds with a pharmaceutial potential is the work carried out for trans-diastereomers of the natural cytokine modulator cytoxazone [84], which also illustrates some of the difficulties associated with the method (eg. that the assignment of the absolute configuration is erroneous when a popular B3LYP functional with a small basis set is used, and that it is essential to account for the presence of a solvent, even if only by means of polarizable continuum model with no explicit solvent molecules).

In principle, the danger of erroneous assigment of the absolute configuration is smaller in the case of vibrational chirooptical spectroscopic methods (VCD or ROA), where there are much more signals than in the electronic spectrum and it is enough if the sign of a majority of them is correctly predicted. However, the use of electronic CD is still more widespread, since the instruments are less expensive and it is possible to use relatively small amounts of the optically pure enantiomer to obtain the spectrum, which is not the case for example of ROA.

The use of electronic CD for determination of absolute configuration is extensively reviewed in Refs. 85 and 86. For the comparison of advantages and disadvantages of using vibrational chiroptical methods for the absolute configuration assignment the reader is referred to Refs. 87 and 88.

6.6 Chirality of Excited States—Circularly Polarized Luminscence

As explained above, ECD is a differential absorption of left- and right circularly polarized light by a chiral molecule. An analogous phenomenon, a differential emission of left- and right circularly polarized light, known as circularly polarized luminscence, CPL, can be observed in fluorescence and phosphorescence spectra. While

ECD brings information on the structure of the ground state of a chiral molecule, CPL is a source of information on the geometry of the excited state, which may differ significantly from the ground-state one. (CPL is not to be confused with fluorescence-detected CD, which still probes the structure of the ground state.) Unlike CD, for which several commercial instruments have been distributed for over 40 years, CPL measurements have been made up to date mainly on custom-built instruments. However, with the method becoming more popular, first commercial spectropolarimeters/fluorometers emerged (the first one has been CPL-200 by JASCO).

Even though CPL is far less widespread than CD, CPL spectra of numerous systems of biological significance have been measured, and it seems this field of application is growing. There are several early works (reviewed in Ref. 89 and 90) concerning CPL as a probe of NADH-dehydrogenase complexes [91, 92], conformational changes in tRNA [93], and bilirubin bound to albumin [94–96], as well as several studies involving time-resolved CPL applied to probe fast processes in biomolecules and their metal complexes (as reviewed for example in Ref. 97). More recently, CPL has also been observed for example in green fluorescent protein [98]. There have been also several studies utilizing strong CPL signal of lanthanide(III) complexes as a structural probe, including the biological applications [99]. More examples of the recent applications of CPL for solving structural problems concerning proteins and their complexes can be found in a specialized review on this subject by Gussakovsky [100] and in a more general review on CPL by Riehl and Muller [101].

Recently, first quantum chemical calculations of CPL have been performed, which is likely to give impulse to more experimental studies. From theoretical point of view, the nature of CPL is very similar to ECD, with the only difference that while ECD intensity needs to be evaluated at the equilibrium structure of the ground state, to calculate CPL intensity one needs to know geometry of the energy minimum on the excited state potential energy surface. This requirement was impossible to fulfill for all but the smallest molecules before developement of analytical excitation energy gradient in time-dependent density functional theory formalism. CPL spectra are being calculated since 2010 [102, 103], and it is likely their calculations will be soon extended to the systems of biological significance.

6.7 SUMMARY

In this brief chapter, we have ventured to outline the most important fields of application of circular dichroism spectroscopy (and related circularly polarized luminescence spectroscopy) in bioscience. While vibrational chiroptical spectroscopies are successful competitors of electronic CD in structural analysis (in particular determination of absolute configuration) of small molecules, ECD is, and is likely to remain, a method of choice in obtaining information on global (and in some cases local, when prosthetic group chromophores are involved) structure of proteins and nucleic acids. Unlike VCD, ECD does not require dissolving biopolymers in D_2O, and, unline in ROA, the samples are not subjected to prolonged laser irradiation.

These are two very important aspects of ECD as a non-degrading method of fast conformational analysis of proteins and nucleic acids. What is more, a short time of a CD measurement makes it a suitable tool for time-dependent studies. Electronic CD and CPL are particularly valuable when the structure of a biomolecule changes in time because of external stimulus (change of solvent, temperature, or complex formation). This makes CD an important tool in drug discovery (when the stimulus is a pharmacologically active molecule), and investigations of such structural changes as protein denaturation, amyloid formation, or DNA melting.

References

1. Biot JB (1817) Mem Acad Roy Sci Inst France 2:41
2. Djerassi C (1960) Optical Rotatory Dispersion. (McGraw-Hill, New York,)
3. Snatzke G (1968) Angew Chem Internat Edit 7:14
4. Haidinger W (1847) Poggendorfs Annalen 70, 531
5. Cotton AM (1896) Ann Chim Physique 8:347
6. Laur P (2012) Comprehensive chiroptical spectroscopy: Volume 1—instrumentation, methodologies, and theoretical simulations. Wiley, Hoboken. (Chap. The first decades after the discovery of CD and ORD by Aime Cotton in 1895)
7. Sreerama N, Woody RW (2000) Circular dichroism: Principles and applications. Wiley, New York, pp 601–620 (Chap. Circular Dichroism of Peptides and Proteins)
8. Cantor C, Schimmel P (1980) Biophysical Chemistry, Part 2: Techniques for the Study of Biological Structure and Function. W.H. Freeman and Company, New York
9. Rodger A, Norden B (1997) Circular Dichroism and Linear Dichroism. Oxford University Press, Oxford. (Oxford Chemistry Masters)
10. Woody RW (2012) Comprehensive chiroptical spectroscopy: Volume 2—applications in stereochemical analysis of synthetic compounds, natural products, and biomolecules. Wiley, Hoboken. (Chap. Electronic Circular Dichroism of Proteins)
11. Toniolo C, Formaggio F, Woody RW (2012) Comprehensive chiroptical spectroscopy: Volume 2—applications in stereochemical analysis of synthetic compounds, natural products, and biomolecules. Wiley, Hoboken. (Chap. Electronic Circular Dichroism of Peptides)
12. Kypr J, Kejnovska I, Bednarova K, Vorlickova M (2012) "Comprehensive chiroptical spectroscopy: Volume 1—applications in stereochemical analysis of synthetic compounds, natural products, and biomolecules. Wiley, Hoboken. (Chap. Electronic Circular Dichroism Spectroscopy of Nucleic Acids)
13. Gray DM (2012) Comprehensive chiroptical spectroscopy: Volume 2—applications in stere- ochemical analysis of synthetic compounds, natural products, and biomolecules. Wiley, Hoboken. (Chap. Circular Dichroism of Protein-Nucleic Acid Interactions)
14. Ellestad GA (2012) Comprehensive chiroptical spectroscopy: Volume 2—applications in stereochemical analysis of synthetic compounds, natural products, and biomolecules. Wiley, Hoboken. (Chap. Drug and Natural Product Binding to Nucleic Acids Analyzed by Electronic Circular Dichroism)
15. Bertucci C, Pistolozzi M (2012) Comprehensive chiroptical spectroscopy: Volume 2—applications in stereochemical analysis of synthetic compounds, natural products, and biomolecules. Wiley, Hoboken. (Chap. Electronic Circular Dichroism in Drug Discovery)
16. Kelly S, Price N (1997) BBA-Protein Struct M 1338:161
17. Bulheller B, Rodger A, Hirst J (2007) Phys Chem Chem Phys 9:2020
18. Kelly S, Jess T, Price N (2005) BBA-Proteins Proteom 1751:119
19. Miles A, Wallace B (2006) Chem Soc Rev 35:39

20. Karabencheva T, Christov C (2010) In: Donev R (ed) in Advances in Protein Chemistry and Structural Biology, Advances in Protein Chemistry and Structural Biology, vol 80, pp 85–115
21. Rosenfeld L (1928) Z Phys 52:161
22. Ruud K, Helgaker T (2002) Chem Phys Lett 352:533
23. Cheeseman JR, Frisch MJ, Devlin FJ, Stephens PJ (2000) J Phys Chem A 104:1039
24. Grimme S (1996) Chem. Phys. Lett. 259, 128
25. Grimme S, Waletzke M (1999) J. Chem. Phys. 111, 5645
26. Autschbach J, Ziegler T (2002) J Chem Phys 116:891
27. Yabana K, Bertsch GF (1999) Phys Rev A 60:1271
28. Pecul M, Ruud K, Rizzo A, Helgaker T (2004) J Phys Chem A 108:4269
29. Skomorowski W, Pecul M, Salek P, Helgaker T (2007) J Chem Phys 127:085102
30. Sebek J, Gyurcsik B, Sebestik J, Kejik Z, Bednarova L, Bour P (2007) J Phys Chem A 111:2750
31. Kaminsky J, Kubelka J, Bour P (2011) J Phys Chem A 115:1734
32. Brkljaca Z, Condic-Jurkic K, Smith A-S, Smith DM J Comp Theor Comp
33. Shcherbin D, Ruud K (2008) Chem Phys 349:234
34. Pecul M, Ruud K, Helgaker T (2004) Chem Phys Lett 388:110
35. Hirst JD (1998) J Chem Phys 109:782
36. Woody RW, Sreerama N (1999) J Chem Phys 111:2844
37. Applequist J, Sundberg KR, Olson ML, Weiss LC (1979) J Chem Phys 70:1240
38. Woody RW (1968) J Chem Phys 49:4797
39. Bayley P, Nielsen E, Schellman J (1969) J Phys Chem 73:228
40. Liu HW, Nakanishi K (1981) J Am Chem Soc 103:5591
41. Liu HW, Nakanishi K (1982) J Am Chem Soc 104:1178
42. Harada N, Nakanishi K (1983) Circular Dichroic Spectroscopy – Exciton Coupling in Organic Stereochemistry. University Science Books, Mill Valley
43. Hirst JD, Hirst DM, Brooks CL (1997) J Phys Chem A 101:4821
44. Besley NA, Hirst JD (1998) J Phys Chem A 102:10791
45. Sebek J, Kejik Z, Bour P (2006) J Phys Chem A 110:4702
46. Rogers D, Hirst J (2004) Chirality 16:234
47. Grishina I, Woody R, (1994) Faraday Discussion 99:245
48. Roy A, Bour P, Keiderling TA (2009) Chirality 21:E163
49. Malon P, Bednarova L, Straka M, Krejci L, Kumprecht L, Kraus T, Kubanova M, Baumruk V (2010) Chirality 22:E47
50. Bednarova L, Bour P, Malon P (2010) Chirality 22:514
51. Robin M (1975) Higher Excited States of Polyatomic Molecules. Academic Press, New York
52. Bulheller B, Miles A, Wallace B, Hirst J (2008) J Phys Chem B 112:1866
53. Bulheller B, Miles A, Wallace B, Hirst J (2008) J Phys Chem B 112:1866
54. Woody R (2010) Chirality 22:E22
55. Zako T, Sakono M, Hashimoto N, Ihara M, Maeda M (2009) Biophys J 96:3331
56. Pedersen J, Andersen C, Otzen D (2010) FEBS J 277:4591
57. Giiler G, Drafic E, Vorobev M, Vogel V, Mantele W (2011) Spectrochim Acta A 79:104
58. Vass E, Majer Z, Kohalmy K, Hollosi M (2010) Chirality 22:762
59. Greenfield N, Fasman G (1969) Biochem 10:41084116
60. Whitmore L, Wallace B (2004) Nucleic Acids Res 32:W668
61. Johnson W (1999) Proteins 35:307
62. Sreerama N, Woody R (2000) Anal Biochem 287:252
63. Louis-Jeune C, Andrade-Navarro M, Perez-Iratxeta C (2012) Proteins 80:374
64. Whitmore L, Wallace B (2008) Biopolym 89:392
65. Abdul-Gader A, Miles A, Wallace B (2011) Bioinforma 27:1630
66. Janes R, Miles A, Woollett B, Klose D, Wallace B (2012) Chirality 24:751
67. Rogers D, Hirst J (2004) Biochem 43:11092
68. Lala A, Kaul P (1992) J Biol Chem 267:19914
69. Dael HV (1993) Biochem 32:11886

70. Demarest S, Boice J, Fairman R, Raleigh D (1999) J Mol Biol 294:213
71. Gawronski J, Grajewski J (2003) Org Lett 5:3301
72. Zsila F (2003) FEBS Lett 539:85
73. Zsila F, Bikadi Z, Simonyi M (2003) Biochem Pharmacol 64:1651
74. Dzwolak W, Pecul M (2005) FEBS Letters 104:6601
75. Dzwolak W, Loksztejn A, Galinska-Rakoczy A, Adachi R, Goto Y, Rupnicki L (2007) J Amer Chem Soc 129:7517
76. Loksztejn A, Dzwolak W (2008) J Mol Biol 379:9
77. Loksztejn A, Dzwolak W (2010) J Mol Biol 395:643
78. Babenko V, Harada T, Yagi H, Goto Y, Kuroda R, Dzwolak W (2011) Chirality 23:638
79. Johnson W (2000) Circular dichroism: Principles and applications. Wiley, New York, pp 703–718. (Chap. Circular Dichroism of Peptides and Proteins)
80. Maurizot J (2000) Circular dichroism: Principles and applications. Wiley, New York, pp 719–739. (Chap. Circular Dichroism of Peptides and Proteins)
81. Gray D (2000) Circular dichroism: Principles and applications. Wiley, New York, pp 769–796. (Chap. Circular Dichroism: Principles and Applications)
82. Kypr J, Kejnovska I, Renciuk D, Vorlickova M (2009) Nucleic Acids Res 37:1713
83. Riazance J, Baase W, Johnson W Jr, Hall K, Cruz P, Tinoco I (1985) Nucleic Acids Res 13:4983
84. Kwit M, Rozwadowska MD, Gawromiski J, Grajewska A (2009) J Org Chem 74:8051, pMID:19817355, http://pubs.acs.org/doi/pdf/10.1021/jo901175s
85. Berova N, Ellestad G, Harada N (2000) Comprehensive Natural Products II Chemistry and Biology, (Elsevier) Chap.Modern Methods in Natural Product Chemistry: Characterization by Circular Dichroism Spectroscopy, pp 91–147
86. Berova N, Bari LD, Pescitelli G (2007) Chem Soc Rev 36:914
87. Stephens PJ, Devlin FJ, Pan J-J (2008) Chirality 20:643
88. He Y ,Wang B, Dukor RK, Nafie LA (2011) Appl Spectrosc 65:699
89. Riehl JP, Richardson FS (1977) Chem Rev 77:773
90. Riehl JP, Richardson FS (1986) Chem Rev 86:1
91. Gafni A (1978) Biochemistry 17, 1301
92. Gafni A (1981) J Biol Chem 256:8875
93. Steinberg N, Wachtel EJ, Gafni A (1982) Biochem 21:2573
94. Tran CD, Drake AF (1981) Biochem Biophys Res Commun 101:76
95. Tran C, Beddard GS (1981) Biochem Biophys Acta 678:293
96. Tran C, Beddard GS (1982) J Am Chem Soc 104:6741
97. Brittain HM (1996) Chirality 8:357
98. Goto H, Sawada I, Nomura N (2010) Int J Pol Mat 59:786
99. Montgomery CP, Murray BS, New EJ, Pal R, Parker D (2009) Accounts of Chemical Research 42:925
100. Gussakovsky E (2010) Reviews in fluorescence 2008. Springer Science + Business Media, Mill Valley, pp 425–459 (Chap. Circularly Polarized Luminescence (CPL) of Proteins and Protein Complexes)
101. Riehl J, Muller G (2012) Comprehensive chiroptical spectroscopy: Volume 1—instrumentation, methodologies, and theoretical simulations. Wiley, Hoboken. (Chap. Circularly Polarized Luminscence Spectroscopy and Emission-Detected Circular Dichroism)
102. Pritchard B, Autschbach J (2010) Chem Phys Chem 11:2409
103. Pecul M, Ruud K (2011) Phys Chem Chem Phys 11:643
104. Nesgaard L, Hoffmann S, Andersen C, Malmendal A, Otzen D (2008) Biopolymers 89:779
105. Sreerama N, Manning M, Powers M, Zhang J, Goldenberg D, Woody R (1999) Biochemistry 38:10814
106. Crescitelli F, Mommaerts W, Shaw T (1966) P Natl Acad Sci USA 56:1729
107. Zsila F, Bikadi Z, Simonyi M (2003) Biochem Pharmacol 65:447
108. Nishii FTI, Kataoka M, Goto Y (1994) Biochemistry 33:4903
109. Babenko V, Dzwolak W (2011) Chem Commun 47:10686

Chapter 7
Biologically Relevant Molecules Studied in Low Temperature Inert Matrices

Rui Fausto and Nihal Kuş

Abstract Matrix isolation is a technique where gaseous atoms and molecules are trapped in an environment of solidified inert gases at temperatures close to absolute zero. The method was originally used to study free radicals and other short-lived chemical reaction intermediates, but receives nowadays many other uses, including its application to the study of molecules of biological importance. Without surprise, the method has been progressively catching the interest of biochemists and molecular biologists, and has been applied to many molecules with pharmaceutical or medical use. In this Chapter, we describe the use of matrix isolation infrared spectroscopy to investigate molecules of biological interest. The fundamentals of the technique are briefly presented, and examples of its application to the study of different phenomena in a series of selected molecules of biological relevance are provided.

Keywords Low temperature inert matrices • Infrared spectroscopy • Amino acids • Nuclei acids' bases • Phenols • Coumarins

7.1 Introduction

Biologically relevant molecules may be large complex molecular entities or simple small molecules. Proteins and nucleic acids are examples of large biomolecules that are fundamental for life, since they constitute the main machinery that makes a living being functioning. However, even the simplest molecules may play a fundamental role in life. Molecular oxygen is, probably, the best example of a simple, small molecule playing a critical role in biology, at least for the nowadays most developed organisms.

Along the biological evolution, the organisms had to use simple, easily accessible molecules as energy-providing materials to keep them working, but they also

R. Fausto (✉) · N. Kuş
Department of Chemistry, University of Coimbra, 3004-535, Coimbra, Portugal
e-mail: rfausto@ci.uc.pt

N. Kuş
Department of Physics, Anadolu University, 26470, Eskişehir, Turkey

M. Baranska (ed.), *Optical Spectroscopy and Computational Methods in Biology and Medicine,* Challenges and Advances in Computational Chemistry and Physics 14,
DOI 10.1007/978-94-007-7832-0_7, © Springer Science+Business Media Dordrecht 2014

used these inexpensive molecular materials as pieces to build up their own bodies and machinery, from controlling and defense systems to structural and reproductive devices. Proteins are made from simple building blocks, the amino acids, which can be produced from non-biomorphic materials under certain circumstances, and that are believed to have been initially formed that way on the Earth. Amino acids also play a critical role in the systems of command of the living beings, as neurotransmitters. The main components of the nucleic acids are relatively simple heterocyclic molecules, belonging to the purine and pyrimidine families, which are the most widely-distributed kind of nitrogen-containing heterocycles in nature [1]. Purines are also significant components in a number of other important biomolecules, such as adenosine- and guanosine- 5'-triphosphates (ATP, GTP), cyclic adenosine monophosphate, nicotinamide-adenine dinucleotide (NAD^+/NADH) and coenzyme A, for example, which are main actors in the fundamental bioprocesses of energy production, intracellular signal transduction, biosynthesis and ribonucleic acid (RNA) transcription. Pyrimidine phosphates participate in the biosynthesis of several molecules too. The life-related purines and pyrimidines can also be obtained from non-biological materials. The parent purine was first synthesized in 1899 from uric acid [2], and purines can nowadays be obtained in the laboratory from formamide [3], the simplest amide. Pyrimidines can also be prepared in the laboratory from very simple molecules, *e.g.*, by reacting amides with carbonitriles under electrophilic activation of the amide [4]. Though the chemistry and biological functioning of proteins and nucleic acids cannot be described by the sum of those of their building blocks, the detailed understanding of the relevant properties of these latter constitutes a fundamental piece of information for improving our understanding of the complex systems and of their way of working.

Along evolution, biological systems had also to develop molecular machinery for recognition of molecules (and other biological systems) that might be either beneficial or dangerous. These recognition systems, such as most of the biological processing and signaling devices (enzymes, prions, nucleic acids, neurotransmitters, hormones, etc), take as essential information at the molecular level the spatial forms of the molecules. Small modifications in a molecule, produced by substituting one of its atoms by a different one or by changing the spatial arrangement of the atoms in the molecule (conformational changes), can lead to dramatic changes in its function or be responsible for the activation of a complex biological process. For example, methylated versions of the major pyrimidine nucleic acid bases (cytosine, thymine, uracil), which can also occur as minor forms in these latter biomolecules, appear to have regulatory functions rather than participating directly in the coding and transcription processes [5]. On the other hand, the importance of molecular conformation in biochemistry is well-recognized. Conformations largely determine, for instance, catalytic mechanisms in enzymes, antifreeze cryoprotectants efficiency and membrane properties, and are the basis of the general phenomenon of molecular recognition. Conformational changes are key events in the vision process [6], in the control of membrane permeability and molecular transport in cells [7], and in

many biosynthetic mechanisms, including that of ATP [8]. Protein folding is also controlled by the conformations of the constituting amino acids.

Ultimately, the energetically-accessible spatial arrangements of the atoms in the molecules (or, in other terms, the conformations assumed by the molecules), determine the physical and chemical properties of substances and their biological functions. The many structural and conformational studies already performed on small molecules constitute nowadays a basis for the understanding of the conformational behavior of the complex biomolecular systems. This fact also stresses the relevance of such studies, in particular when they focus on molecules that are building blocks of biological macromolecules (*e.g.*, amino acids, nucleic acids bases) and can be performed under precisely controlled conditions, such as those achievable for a molecule isolated in a low temperature inert matrix.

7.2 The Matrix Isolation Method

Matrix isolation is a technique where gaseous atoms and molecules are trapped in an environment of solidified inert gases at temperatures close to absolute zero. The method was originally used to study free radicals and other short-lived chemical reaction intermediates, but receives nowadays many other uses. By combining this trapping technique with different spectroscopic detection methods a powerful research tool has evolved, which is being applied in many laboratories worldwide in a large number of research areas.

The matrix isolation method was first developed almost simultaneously by the groups of Pimentel, at Berkeley (USA), and Porter, at Cambridge (UK) [9, 10]. The name of the method was coined in the historical paper by Pimentel, Whittle and Dows "*Matrix Isolation Method for the Experimental Study of Unstable Species*", a single-page paper which appeared in the Journal of Chemical Physics, in 1954 [9]. Very interestingly, matrix isolation only started to become popular among organic chemists almost 20 years later, after publication of 3 papers, nearly at the same time, by three different groups, where the experimental observation of the hitherto elusive cyclobutadiene was reported. Cyclobutadiene was produced photochemically *in situ* in inert argon matrices containing the precursor α-pyrone [11–13]. α-Pyrone was found to undergo a fast photoequilibration with its isomeric open-ring aldehyde-ketene and a much slower ring closure to a bicyclo β-lactone (the α-pyrone Dewar isomer), which in turn eliminated CO_2 to give cyclobutadiene (Fig. 7.1). As mentioned above, at that time cyclobutadiene had never been experimentally characterized, and these studies received great impact among the scientific community.

Though matrix isolation was invented as a tool for studies of reactive species, the method has also some advantages over other techniques for the study of stable molecules, and has been used extensively for this purpose too. In general, matrix isolation is useful when we are interested (i) to observe directly and characterize

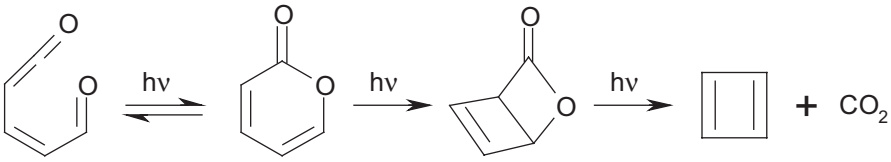


Fig. 7.1 Photochemical reaction pathways observed for matrix-isolated α-pyrone: fast photo-equilibration with the conjugated aldehyde-ketene and slow ring-closure to the Dewar isomer. The Dewar isomer of α-pyrone subsequently expels CO_2, to produce cyclobutadiene [11–14]

reaction intermediates or novel reactive species that can be produced *in situ* (for instance, by photolysis of a suitable precursor) or in the vapor phase during deposition, (ii) to study molecular species that easily aggregate, isomerize or decompose under other experimental conditions and (iii) to freeze out and study particular molecular conformations or isomeric structures.

Among the main advantages of the method, we can highlight the following, which are consequences of the low working temperature, usually chosen low concentration, rigidity of the matrix and chemical inertness of environment: (a) solute-solute interactions are practically absent in diluted matrices, and can be easily controlled by varying the matrix concentration or annealing the matrix to a higher temperature; (b) solute-solvent interactions can be most of times neglected for many practical applications, since inert gases are generally used as matrix hosts. In fact, one can control the importance of the solute-solvent interactions by proper selection of the matrix host gas, or even use a reactive matrix gas (*e.g.*, oxygen), if interested to take advantage of this; (c) molecular diffusion and rotation are inhibited (except in the case of very small molecules). Because of this, the vibrational spectra are essentially pure vibrational spectra, enabling a direct and easy comparison with theoretically predicted spectroscopic data. In addition, if fragmentation of a matrix-isolated species takes place, most of times the obtained fragments stay in the matrix site where they were formed. Then, no subsequent cross-reactions involving species resulting from fragmentation of different reactant molecules can occur, strongly reducing the number of possible products in comparison with gas phase or solution studies; (d) hot vibrational transitions as well as subtractive combination tones are suppressed, since excited vibrational states are depopulated at the low work temperatures typical of the matrix isolation experiments, introducing an additional simplification in the spectra and contributing to their increased resolution.

In addition to the above-mentioned main advantages, standard matrix isolation studies do not require expensive equipment (Fig. 7.2) and practically all specific instrumentation for matrix preparation can be easily manufactured or transformed locally, once a moderately well equipped machine shop exists in the research institution.

In spite of its great capabilities, matrix isolation IR spectroscopy does also have some drawbacks. Firstly, the studied compound must be evaporated during

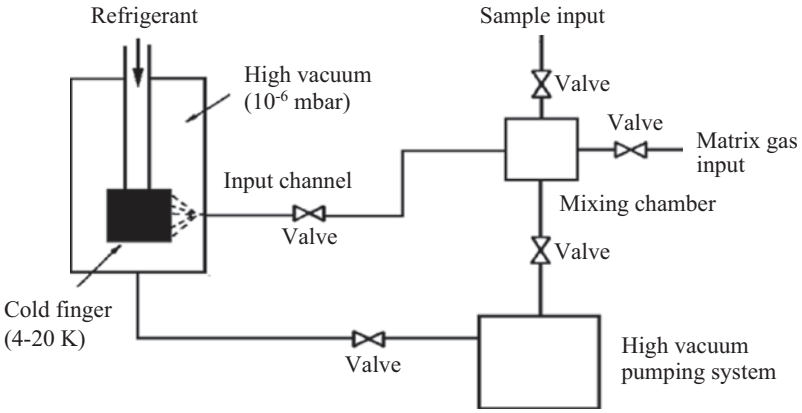


Fig. 7.2 Schematic representation of a matrix-isolation basic set up. The cryostat shall have several external windows in order to enable to register the spectra and perform *in situ* irradiation of the sample whenever required. Different techniques for matrix preparation have been developed, from pulsed methods to continuous flux methods [15–17]. In the first case, pre-mixing of the compound to be studied with the matrix gas (in the picture) is in general necessary, while in the second both pre-mixing or co-deposition of solute and solvent can be used

preparation of the sample, without decomposing. This requirement precludes matrix isolation studies of strongly non-volatile species or thermolabile molecules, and constitutes the most important restriction to its use in the study of complex biomolecules. Secondly, the quantitative determination of matrix concentrations is not an easy task, even if precision low temperature quartz microbalance is used. Hence, matrix isolation spectroscopy has not been used frequently in quantitative analysis.

In third place, it is well known that even under well-controlled deposition conditions the produced matrices allow for different local environments around the solute molecules. The co-existence of these different matrix sites leads to slightly different spectral signatures of molecules of the solute, which results in the usual observation of bands with multiple structure in the spectra of matrix-isolated species. Such problem can, however, be in general minimized by use of different matrix gases and judicious choice of the temperature of the cold substrate during deposition. Finally, since one of the main advantages of the matrix isolation method is the high spectral resolution it allows to achieve, when the method is used to study very large, complex molecules exhibiting extensive intrinsic band overlapping, it loses relevance. This is also an important shortcoming for the general application of the method in biochemistry. Nevertheless, many bioactive or biologically essential compounds of moderate size have been studied by matrix isolation infrared spectroscopy, and the method has been progressively catching the interest of biochemists as well. This is the case of the amino acids, DNA bases, phenolic compounds and coumarins, which will be shortly reviewed in the next sections of this Chapter mostly based on our own studies on these types of compounds, but also of many other molecules with pharmaceutical or medical use [18–20].

7.3 Amino Acids

In the crystalline state, amino acids exist as zwitterions. Upon sublimation, however, these forms convert into the non-ionic forms. Very interestingly, fast deposition of amino acid vapors onto a cold (~10 K) infrared-transparent substrate allowed spectroscopic observation of the non-ionic forms of glycine, sarcosine and *N,N*-dimethylglycine in their neat amorphous solid phases [21]. The spectra obtained immediately after deposition revealed the presence in the initially produced amorphous film of both the non-ionic and zwitterionic forms of the studied amino acids. The identification of the non-ionic forms and their vibrational signatures could be unequivocally established by taking into account data for the matrix-isolated non-ionic monomeric and aggregated amino acids [22–24] as well as solid state data for acid and alkaline salts of glycine [25]. Upon temperature increase, the non-ionic forms were found to convert non-reversibly into the zwitterionic forms (Fig. 7.3). Observation of aggregates of non-ionic forms of the amino acids in the solid phase, together with the fact that the amount of these forms diminished with increasing both the temperatures of the cold substrate and of the vapor prior to deposition, enabled to conclude that the observed (non-ionic form)→zwitterion conversion occurs in the solid state, where the zwitterionic species is most stable, and that intermolecular interactions in this phase play an important role in the mechanism of proton transfer [21].

Detailed structural and spectroscopic characterizations of the monomeric non-ionic forms of many simple amino acids have been achieved by combining matrix isolation IR spectroscopy with contemporary methods of quantum chemistry.

The simultaneous experimental observation of the three lower energy conformers of the simplest amino acid, glycine, by matrix isolation infrared spectroscopy [22], has been a milestone in the structural characterization of this compound. Previous microwave spectroscopy studies [26–30] had already provided experimental evidence of existence of two conformers of glycine in the gas phase (I and II; Fig. 7.4), though the significantly lower dipole moment of the most stable conformer I (bearing an intramolecular hydrogen-bond of the $NH_2\cdots O{=}C$ type), compared to conformer II (possessing an intramolecular O–H···N hydrogen-bond), initially made detection of the first one difficult by that technique. Conformer III (with an $NH_2\cdots$OH H-bond; Fig. 7.4) has also a low dipole moment, and its expected small relative population makes its experimental detection by microwave spectroscopy very hard. On the other hand, a gas phase electron diffraction study [31], had suggested that, at 219 °C, about 24 % of the compound should correspond to a mixture of minor conformers of glycine, which were supposed to be conformers II and III. However, these conformers are different from each other only in the position of the hydrogens, and they are indistinguishable by their electron scattering intensities. Thus, the question of the conformational structure of glycine in the gas phase could only be effectively solved after its study by matrix isolation infrared spectroscopy. In the matrix isolation study [22], the compound and some of its deuterated isotopologues were isolated in low-temperature argon matrices, and the infrared

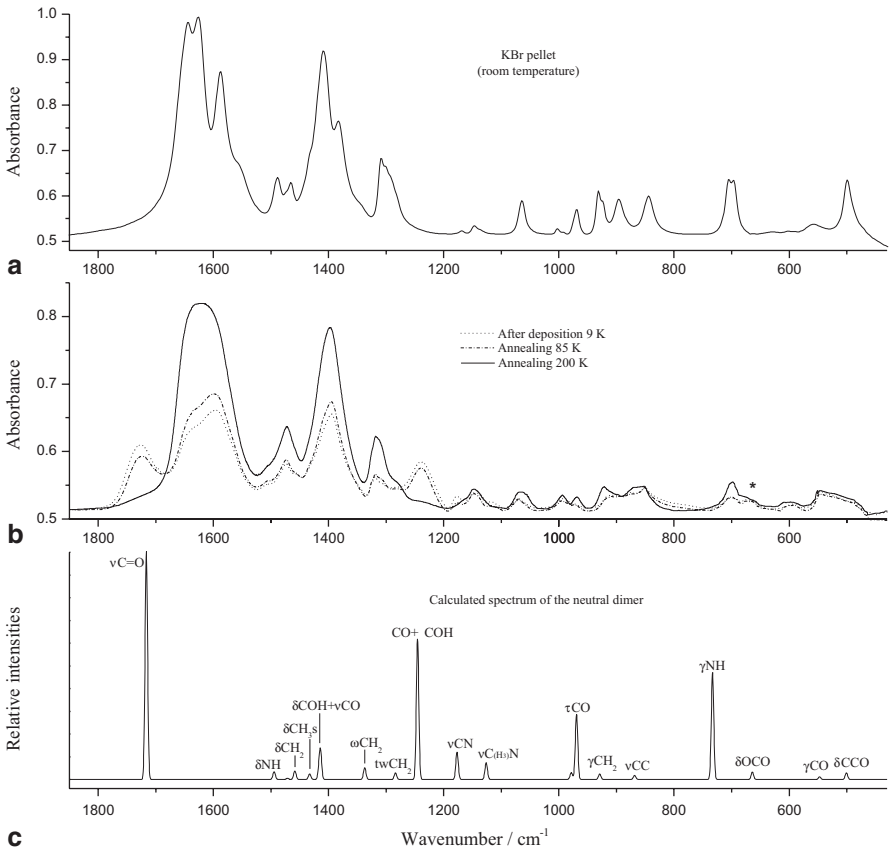


Fig. 7.3 **a** Rom temperature IR spectrum of crystalline sarcosine in a KBr pellet; **b** annealing experiments carried out on the sarcosine low temperature solid film; **c** DFT(B3LYP)/6-311++G** calculated IR spectrum of the centrosymmetric carboxylic acid dimer of non-ionic sarcosine, with monomeric units in their most stable conformation [23]. (Reproduced from [21] with permission from the PCCP Owner Societies)

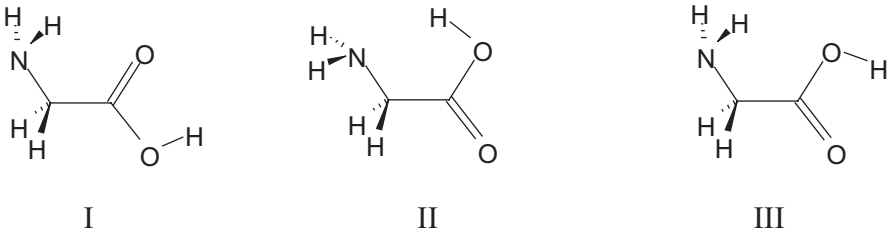


Fig. 7.4 – The lowest energy conformers of non-ionic glycine, I, II and III

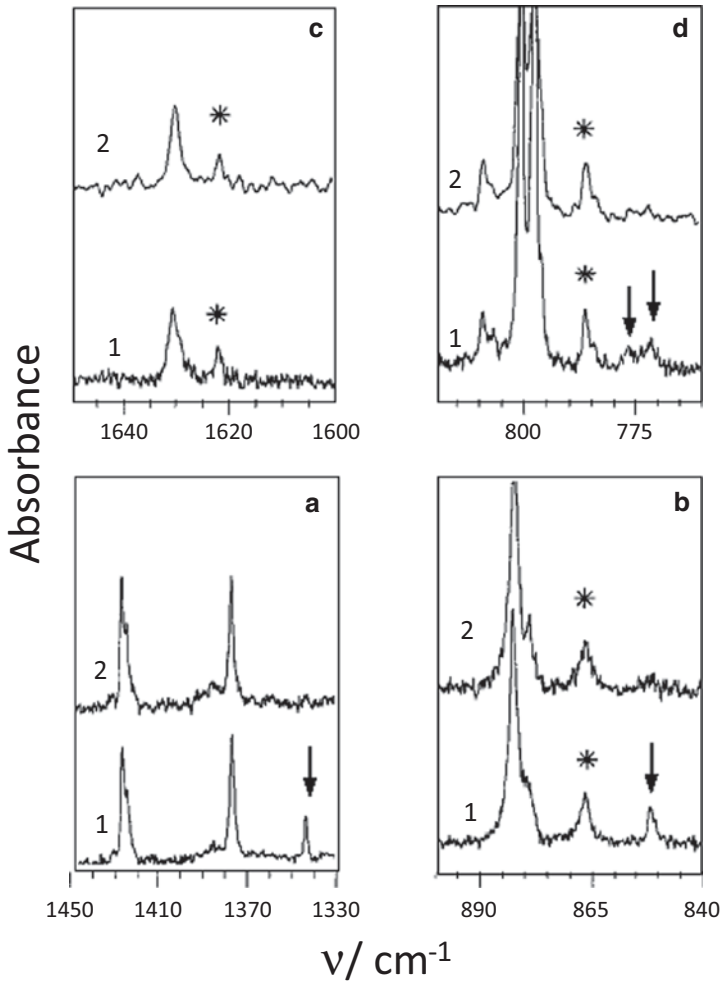


Fig. 7.5 IR spectral regions of the glycine. OH bending/C-O stretching vibration region (**a**), CCOH torsion/NH_2 bending vibration region (**b**), NH_2 scissors vibration region (**c**), NH_2 bending vibration region (**d**). (*1*) IR spectrum recorded after matrix deposition at 13 K. (*2*) IR spectrum recorded after matrix annealing at 20 K for 30 min. Asterisks (*) mark the bands assigned to conformer II. Down arrows (↓) mark the bands assigned to conformer III. (Adapted with permission from [22], copyright © 1998 American Chemical Society)

spectral characteristics of the three most stable conformers could be determined and assigned. Conversion of conformer III into the most stable conformer I was observed upon annealing of the matrices at temperatures higher than 13 K, while the population of conformer II was not affected by the increase of the temperature of the matrix (Fig. 7.5). Band intensity changes with temperature thus allowed, for the first time, the unequivocal experimental identification of the three lowest energy conformers of glycine.

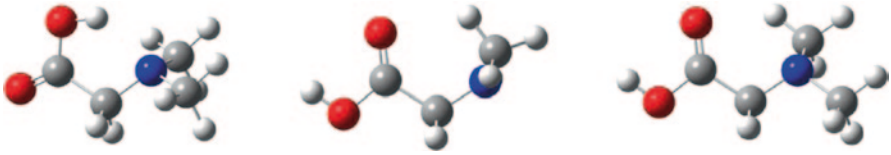


Fig. 7.6 Lowest energy conformers of non-ionic DMG

Matrix isolation IR spectroscopic studies also provided the first experimental observation and spectral signature characterization of the three lowest energy conformers of *N,N*-dimethylglycine (DMG) [24]. Detailed analysis of the spectra obtained in both argon and xenon matrices for the parent (O–H) DMG and its O–D isotopologue demonstrated that, for this molecule, the conformational ground state is the intramolecularly O–H···N hydrogen-bonded GAT conformer (Fig. 7.6), where the Lp-N–C–C (Lp = nitrogen lone electron pair) and N–C–C = O dihedral angles are 30° and *ca.* 180°, respectively, and the carboxylic group assumes the *trans* configuration. In the other two experimentally relevant forms (ASC and GSC; Fig. 7.6), the carboxylic moiety assumes the *cis* conformation, while the N–C–C = O axis adopts the *syn* arrangement and the conformation around the N–C bond is *anti* and *gauche*, respectively.

The energy barriers for conformational interconversion between the three experimentally observed conformers of DMG were found to be considerably large [24], thus preventing the occurrence of these processes in the matrices. However, very interestingly, in the xenon matrix, aggregation of DMG was found to be dependent on the initial conformational state of the monomers, the ASC conformer exhibiting a considerably higher ability to aggregate than both the GAT and GSC forms [24]. It is not surprising that the GAT conformer, in which both OH and amino groups are involved in a strong intramolecular hydrogen bond, is less able to aggregate than the ASC and GSC forms. On the other hand, *a priori*, the relative ability of the two higher energy conformers to aggregate could not be expected to be markedly different, since both forms have the relevant groups available for participation in intermolecular hydrogen bonding interactions. However, the observed experimental trends could be explained in terms of several structural features of the two conformers favoring the easier aggregation of conformer ASC: larger positive charge on the hydroxyl hydrogen and larger negative charge on the carbonyl oxygen; higher symmetry (leading to smaller matrix packing perturbation); more accessible nitrogen lone electron pair; and greater relative stability of the ASC dimer compared to those established by GSC [24]. To the best of our knowledge, this was the first reported observation of a conformationally dependent aggregation trend. Such phenomenon has been described since then for other compounds [32, 33] and can in fact expected to be relevant in biochemical terms.

In case of sarcosine (*N*-methylglycine), matrix isolation infrared studies [23] revealed that the most stable conformer, ASC (Fig. 7.7), is similar to the glycine most

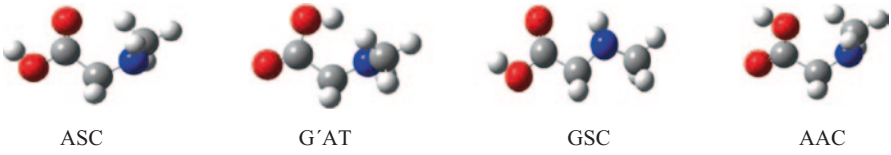


Fig. 7.7 Sarcosine conformers observed in low temperature matrices

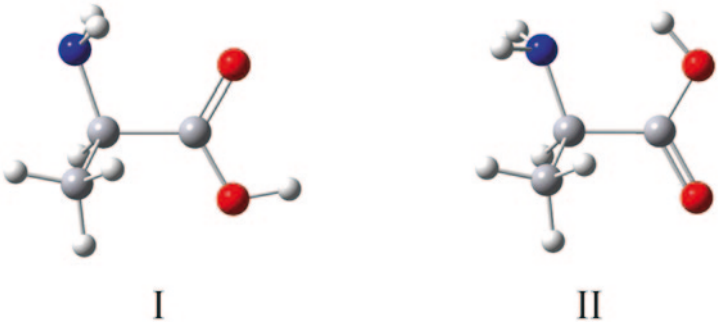


Fig. 7.8 The two low-energy conformers of non-ionic α-alanine observed in cryomatrices

stable conformer and is characterized by an intramolecular N–H···O=hydrogen bond and a *cis* carboxylic group, while the second most stable conformer (G'AT) exhibits a strong O–H···N intramolecular hydrogen bond and a *trans* carboxylic group, being similar to the most stable form of *N,N*-dimethylglycine. In consonance with theoretical predictions [23], the infrared spectra of the matrix-isolated compound revealed the presence of four conformers, ASC, G'AT, GSC and AAC (see Fig. 7.7), with the two most stable forms accounting for *ca.* 70% of the total population in the gas phase prior to deposition.

Matrix-isolated α-alanine was found to exist in two different conformers (Fig. 7.8) [34]. Like in the case of glycine, the most stable conformer of α-alanine (conformer I) bears an intramolecular hydrogen-bond of the NH_2···O=C type, while the second experimentally relevant conformer (II) possesses an O–H···N intramolecular H-bond. These structural features were clearly evidenced in the matrix isolation infrared spectra of the compound. The most stable form was found to contribute to ~90% of the total conformational population in the as-deposited matrices [34]. On the other hand, some α-alanine conformers with predicted low relative energies were not found to be present in the matrices, what was attributed to the existence of low-energy barriers between these conformers and the two observed forms. The non-observed higher-energy conformers can then be converted into the most stable forms during matrix deposition (conformational cooling) [34].

Compared to glycine, α-alanine and sarcosine, the relative order of stability of the two types of low-energy conformers usually found in simple amino acids was found to be reversed in the case of proline [35], where the lowest energy conformer

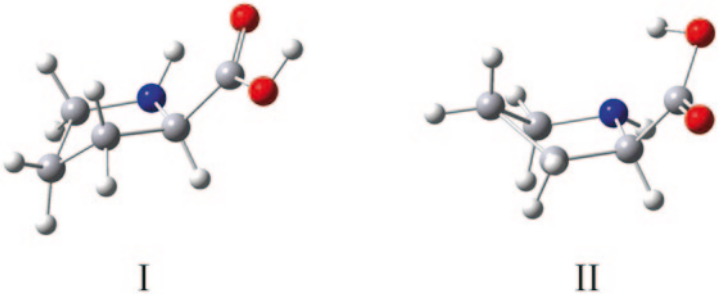


Fig. 7.9 The two low-energy conformers of non-ionic proline observed in cryomatrices

correspond to an O–H···N hydrogen-bonded form (conformer II; Fig. 7.9) and the second observed conformer to a N–H···O=hydrogen-bonded species (conformer I). It was shown that the intramolecular O–H···N H-bond in the proline conformer II is considerably stronger than for other amino acids, since the observed red shift of the OH stretching vibration due to this interaction is unusually larger (534 cm^{-1}, *vs.* 340–360 cm^{-1} in other amino acids) [21–24, 34, 35].

Serine was found to possess a considerably more complex conformational space than the above mentioned amino acids [36, 37]. A systematic investigation of the conformational potential energy surface of non-ionic serine revealed the existence of 61 different minima [36]. Those with relative energies within 7 kJ mol^{-1} were estimated to account for *ca.* 93% of the total conformational population. In low temperature argon matrices, different conformers were observed [37], which can be classified accordingly to the main intramolecular interaction they exhibit: $O–H_A···N$, $O–H_C···N$ or $O–H_A···O$=hydrogen bonds (where the subscripts A and C stand for alcohol and carboxylic group, respectively). The $O–H_C···N$ intramolecular interaction was found to be considerably stronger than both the $O–H_A···N$ and $O–H_A···O$=hydrogen bonds, and leads to reduce the abundance of conformers bearing that interaction at high temperatures, due to entropy effects. Similar entropic effects were also found to exist in phenylalanine [38], where the conformers having a weaker H-bond of the N–H···O=type, being considerably more flexible than the lowest energy forms possessing the much stronger intramolecular H-bond of the O–H···N type, are stabilized by entropy. At 423 K, the entropically favored conformers were found to be the first and third most abundant species present in the conformational equilibrium, with relative populations of *ca.* 15% each, whereas their populations could be expected to be only *ca.* 5% if entropy effects were not taken into consideration [38]. Phenylalanine can indeed be considered a notable example of a molecule where entropy plays an essential role in determining the relative abundance of the low-energy conformational states, and then the thermodynamics of the compound, even at moderate temperatures. Entropy effects were also found to be important in determining the relative populations of the different conformers of tryptophan in gas phase [33].

Six phenylalanine conformers were suggested to contribute to the spectrum of the isolated compound (Fig. 7.10), with several low-energy conformers being converted to these six forms during matrix deposition [38]. In the case of tryptophan, analysis

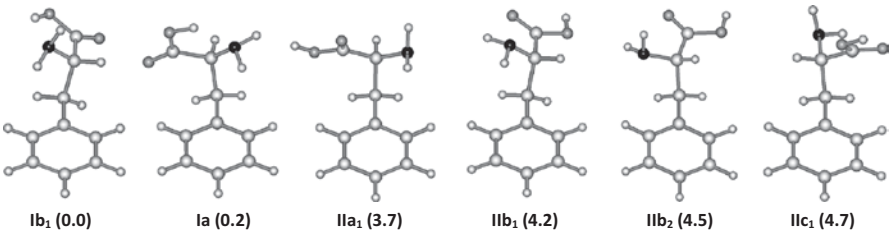


Fig. 7.10 The 6 conformers of non-ionic phenylalanine observed in cryomatrices (naming as in [38])

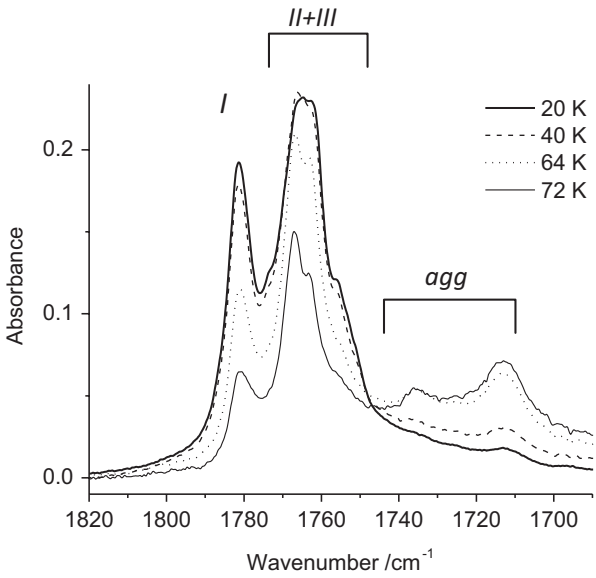


Fig. 7.11 Carbonyl stretching region of the IR spectra (as-deposited and annealed) of tryptophan isolated in solid xenon. *I-III* and *agg* denote absorptions associated with conformers of type I-III and aggregates of the compound, respectively. (Adapted with permission from [33], copyright © 2007 American Chemical Society)

of prominent spectral features unequivocally demonstrates the presence of at least two different types of conformers, the O–H···N hydrogen-bonded forms exhibiting the *trans* arrangement of the carboxylic group (type I), and type II/III conformers, featuring the N–H···O=H-bond and the *cis* conformation of the carboxylic group [33]. Annealing of the xenon matrices resulted in conformer-selective aggregation of tryptophan, with conformers of type I aggregating easier than type II conformers (Fig. 7.11 [33]). In the conformers of type I, both the C=O group and the protons of the NH_2 group are not involved in any intramolecular H-bond, so that they can aggregate easier than the other conformers bearing N–H···O=intramolecular H-bonds. Due to the low overall abundance of conformers of type III, no conclusions could be extracted in relation with their relative ability to aggregate, but it was suggested that these conformers should follow the same trend as type II conformers [33].

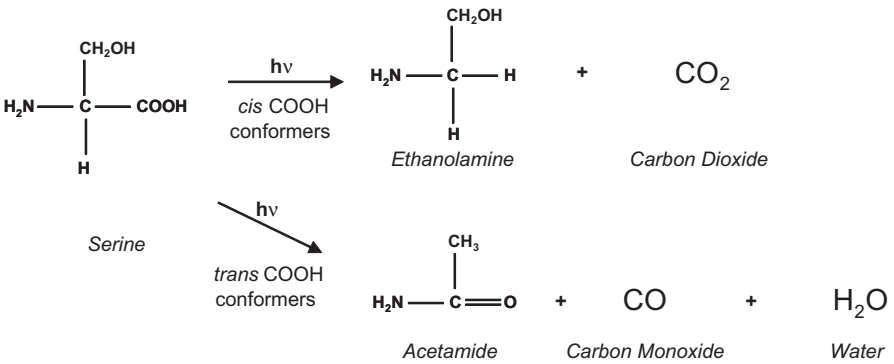


Fig. 7.12 Unimolecular photodecomposition channels observed in the UV ($\lambda > 200$ nm) photochemistry of matrix-isolated serine. (Adapted from [37] with permission from Elsevier)

In tryptophan, besides intramolecular hydrogen-bonding interactions, specific intramolecular electrostatic interactions also significantly contribute to determine the relative stability of the different conformers, with stabilizing interactions of this type involving the pyrrole ring being considerably more important that those involving the phenyl ring of the molecule [33]. Such trend seems to have biochemical relevance, since it has also been noticed for tryptophan residues in proteins [39].

Both serine and phenylalanine were also investigated by matrix isolation infrared spectroscopy in relation to their photochemistry [36–38]. When submitted to UV irradiation ($\lambda > \sim 200$ nm), serine decomposes by two different channels in the S_1 excited state surface: decarboxylation, leading to production of CO_2 and ethanolamine, and decarbonylation, with formation of CO, H_2O and acetamide (Fig. 7.12). Very interestingly, the two observed photoprocesses were found to be dependent on the conformation assumed by the reactant molecule, with conformers bearing $O–H_A\cdots N$ or $O–H_A\cdots O$=hydrogen bonds undergoing decarboxylation and those possessing an $O–H_C\cdots N$ H-bond decarbonylating [36, 37]. This was one of the first observations of conformationally dependent photochemical processes, a nowadays very active field of research. The results also suggest that the transition states' geometries of the carboxylic group for the two processes in the S_1 surface should resemble the *cis* and *trans* arrangements, respectively. For matrix-isolated phenylalanine, the decarboxylation, leading to production of CO_2 and phenethylamine, was the sole observed photochemical pathway [38].

7.4 Nucleic Acid Bases

Among the nucleic acid bases, the most investigated one has been uracil, due to its simpler structure. In solution, at pH = 7, uracil exists as a mixture of two tautomers, amide and imide (Fig. 7.13). On the other hand, in crystalline phase [40], as well

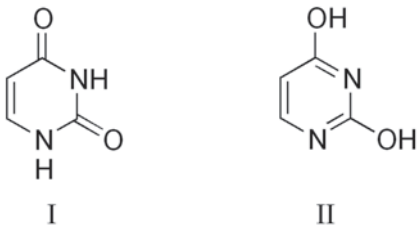


Fig. 7.13 Amide (*I*) and imide (*II*) tautomers of uracil

as both in the gas phase and in cryogenic matrices prepared from the vapor of the compound, uracil exists in a single tautomeric form (amide) [41–48].

The detailed assignment of the infrared spectra of uracil in argon and nitrogen matrices was first reported *ca.* 30 years ago [41, 42]. An interesting feature of the spectra of matrix-isolated uracil is the extensive presence of band-splitting due to Fermi resonances.

UV-photolysis of uracil in an argon matrix, yielded isocyanic acid, most probably formed together with acetylene, though no bands due to this latter species were experimentally observed [49]. The study of the photochemistry of the nucleic acid bases is of interest because genetic mutations may be caused by UV radiation exposition. It is well-known that nucleic acid bases absorb UV radiation efficiently, though when associated in DNA forming hydrogen-bonded pairs, fast non-radiative dissipation of the energy gained by these molecules upon UV absorption makes DNA highly photostable [50, 51]. Uracil and adenine were, however, shown to be considerably photolabile [49] and cannot survive in environments where a significant UV flux is present. Nevertheless, it has still been proposed that nucleic acid bases molecules might first have been produced in the space, since traces of nucleobases have been found in carbonaceous meteorites [49, 52–55].

Thymine (5-methyluracil) also exists in a single tautomeric form (amide) in gas phase and matrices [45, 46, 56, 57], whereas in solution, like uracil, thymine can be found in both amide and imide forms. On the other hand, several tautomeric forms of cytosine have already been experimentally observed, as described below in detail. Among the nucleic acid bases derived from pyrimidine, cytosine has been shown to be the most challenging molecule for structural investigation, and matrix isolation has been the key-technique in the structural studies of this compound.

In case of the purines, adenine revealed itself as a simple case, while guanine was found to be a more complex system from the structural point of view. In solution, both compounds may exist as a complex mixture of several tautomers. However, in case of adenine, though it has been initially believed that the compound existed in two amino tautomeric forms nearly equally populated in the gas phase [58–60] (forms I and II in Fig. 7.14), it was latter on unequivocally demonstrated that only one tautomer (I) is significantly populated in that phase [61, 62]. Isotopic labeling (^{15}N) experiments undertaken on matrix-isolated adenine, together with results from quantum chemical calculations, were essential to definitely solve that question [62].

On the other hand, at least three tautomeric forms of guanine were spectroscopically identified in cryogenic matrices (the amino-oxo forms I and II, and the

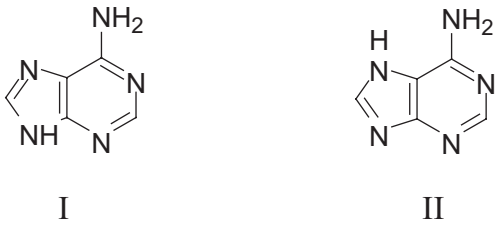


Fig. 7.14 Adenine tautomeric forms. Tautomer *I*, N(9)H, is the only tautomer experimentally observed for the matrix-isolated compound; tautomer *II*, N(7)H, was initially supposed also to be present, but later on excluded

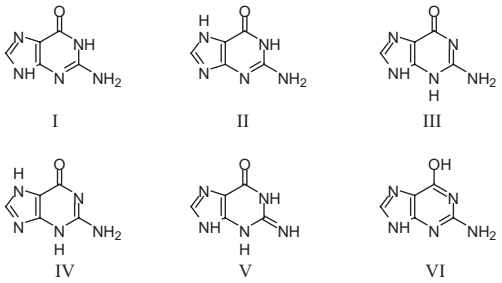


Fig. 7.15 Tautomeric forms of guanine. Tautomers *I*, *II* and *VI* were shown to be present in gas phase and as-deposited cryomatrices of the compound

amino-hydroxy form VI represented in Fig. 7.15) [63–65]. Study of tautomeric species in nucleic acid bases is important because rare tautomers of these compounds can be related to point mutations during DNA duplication [66–69].

Regarding tautomerism and conformational isomerism, cytosine is by far the most interesting nucleic acid basis. Recently, five different forms of cytosine have been observed experimentally [70]. Results of theoretical calculations [71–75] have predicted that the most stable tautomer of cytosine in the gas phase should be the amino–hydroxy form. This tautomer may exist in two conformers (AHI and AH2 in Fig. 7.16) differing by rotation of the hydroxyl group by *ca.* 180°, with AH1 being predicted to be more stable than AH2 by 2.9–3.1 kJ mol^{-1}. The amino–oxo form (AO in Fig. 7.16) was predicted to have a higher energy than AH1 by 4.9–6.3 kJ mol^{-1}, while the two imino–oxo conformers have relative energies to the most stable species equal to *ca.* 7 (OI1) and 12 (OI2) kJ mol^{-1}. According to the calculations, together with the dominating AH1 and AH2 isomers, also AO and IO1 forms should be significantly populated in the gas phase, whereas the relative population of IO2 should be considerably lower, and all the remaining possible forms of cytosine (whose relative energies are much higher) should not be populated at all. Microwave and photoelectron spectroscopic studies confirmed the presence in gas phase of forms AH1, AH2 and AO [76, 77], but the experimental simultaneous observation of all the five lower energy forms of cytosine was only achieved by matrix isolation infrared spectroscopy [70].

The identification of the different species in the matrices was made possible because an elegant strategy for optically inducing isomerization reactions for matrix-isolated cytosine was used [70, 78]. In a first experiment, monomers of cytosine isolated in an argon matrix were irradiated using a narrowband, tunable

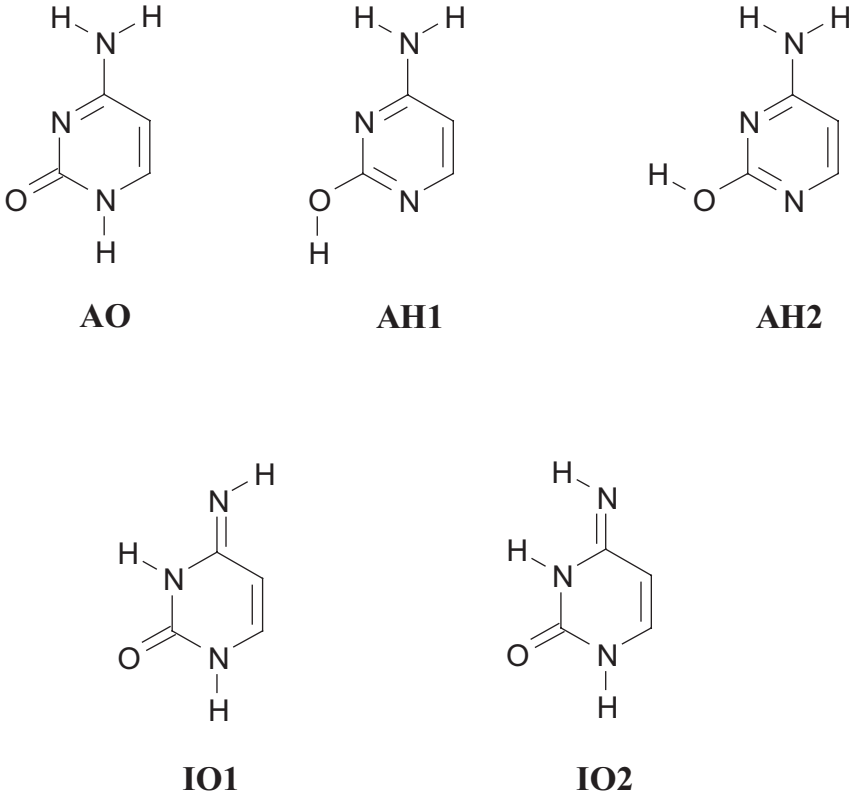


Fig. 7.16 Lowest energy isomeric forms of cytosine

near-IR light emitted by a Quanta-Ray MOPO-SL pulsed optical parametric oscillator at different wavelengths, corresponding to the overtone absorptions of the two amino-hydroxy forms. Upon irradiation at 7,013 cm^{-1}, the initially most populated AH1 conformer converted into AH2, whereas upon irradiation at 7,034 cm^{-1} the reversed reaction was observed. The changes were monitored by observation of the mid-IR spectra of the irradiated matrix (Fig. 7.17) [78].

Use of narrowband selective near-IR irradiation of matrix-isolated small molecules, in particular small carboxylic acids, to induce conformational changes had already been successful [79–81]. However, before the study on cytosine, these phenomena had never been reported for larger heterocyclic molecules. In polyatomic molecules containing more than ten atoms, the higher density of vibrational states may affect the pathways of vibrational energy redistribution and, hence, the efficiency of rotational isomerization stimulated by excitation of the overtones of stretching vibrations. In this regard, the study on cytosine [78] has also constituted a noteworthy experimental achievement.

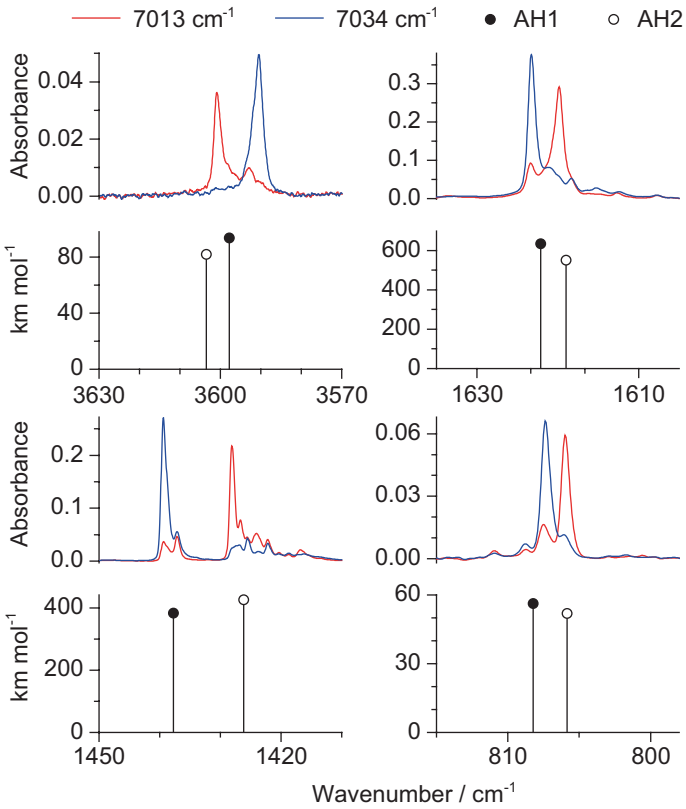


Fig. 7.17 Fragments of the IR spectra of cytosine isolated in an Ar matrix: (*red*) after irradiation at 7,013 cm^{-1}; (*blue*) after irradiation at 7,034 cm^{-1}; compared to the (scaled) IR spectra calculated for *AH1* and *AH2* conformers at the DFT(B3LYP)/6-31++G(d, p) level. (Reproduced from [78] with permission from the PCCP Owner Societies)

Very interestingly, it was recently demonstrated that, upon broadband near-IR/mid-IR irradiation (with the IR source of a FTIR spectrometer) the population ratio of the two amino-hydroxy conformers of cytosine changes towards a ratio corresponding to a photostationary state [82]. Evolution of the conformer population ratio towards the photostationary ratio occurs independent of the initial ratio of conformers, which could be prepared by a population shift (in favor of one of the forms) induced by narrowband selective near-IR excitation (Fig. 7.18). In addition, spontaneous tunneling conversion of the higher-energy AH2 conformer into the lowest-energy AH1 form was also observed for cytosine isolated in a low-temperature argon matrix kept in the dark. This process was found to be slow, occurring on a time scale of days, and, due to the inhomogeneous nature of the matrix media, follows a dispersive type of kinetics rather than the classical mono-exponential kinetics (Fig. 7.19) [82].

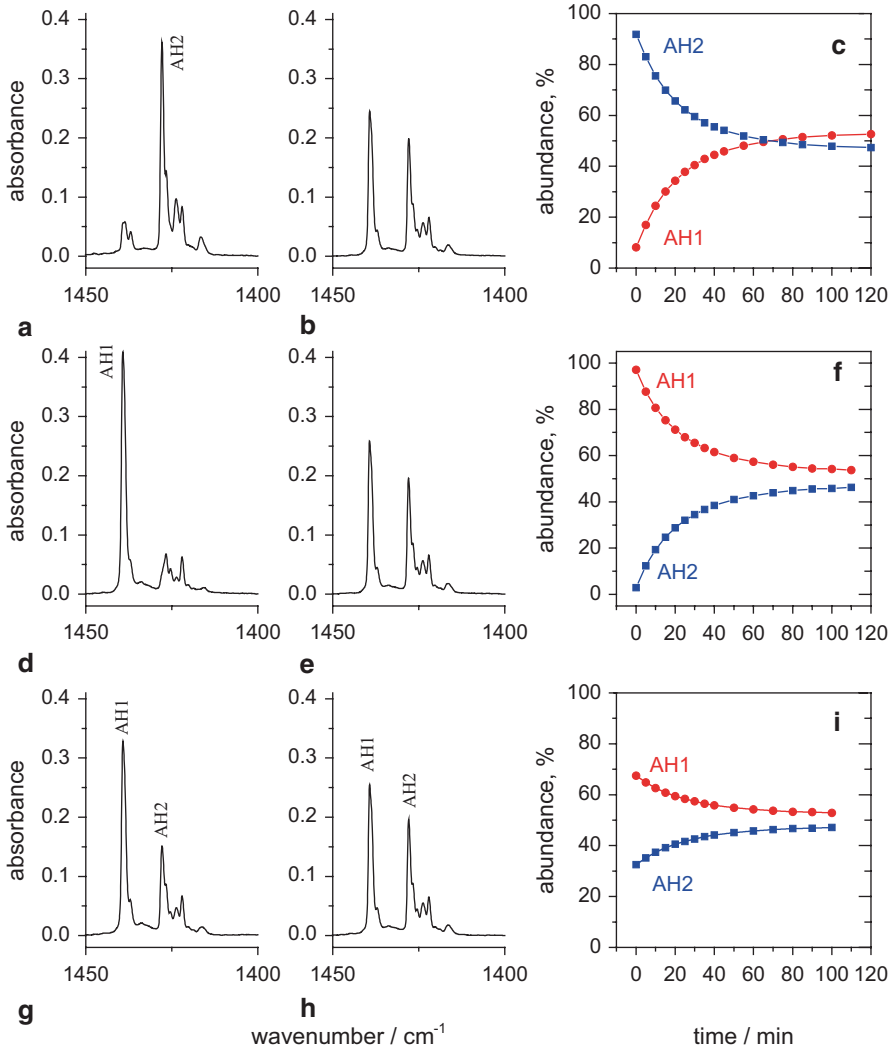


Fig. 7.18 Fragment of the infrared spectrum of cytosine isolated in an Ar matrix: **a** recorded after narrowband near-IR irradiation at 7,013 cm^{-1}. **b** after subsequent 120 min. of exposure to the near-IR/IR broadband radiation of the spectrometer source. **c** evolution of abundances of *AH1* and *AH2* rotamers with time of broadband near-IR/IR irradiation [initial point corresponds to (**a**), final point corresponds to (**b**)]. **d** recorded after narrowband near-IR irradiation at 7,034 cm^{-1}. **e** after subsequent 110 min. of exposure to the near-IR/IR broadband radiation of the spectrometer source. **f** evolution of abundances of *AH1* and *AH2* rotamers with time of broadband near-IR/IR irradiation [initial point corresponds to (**d**), final point corresponds to (**f**)]; **g** recorded after deposition of the matrix, monitored only through a filter transmitting in the spectral range 1,750–800 cm^{-1}; **h** after subsequent 100 min. of exposure to the near-IR/IR broadband radiation of the spectrometer source; **i** evolution of abundances of *AH1* and *AH2* rotamers with time of broadband near-IR/IR irradiation [initial point corresponds to (**g**), final point corresponds to (**h**)]. (Reproduced with permission from [82], copyright © 2012 American Institute of Physics)

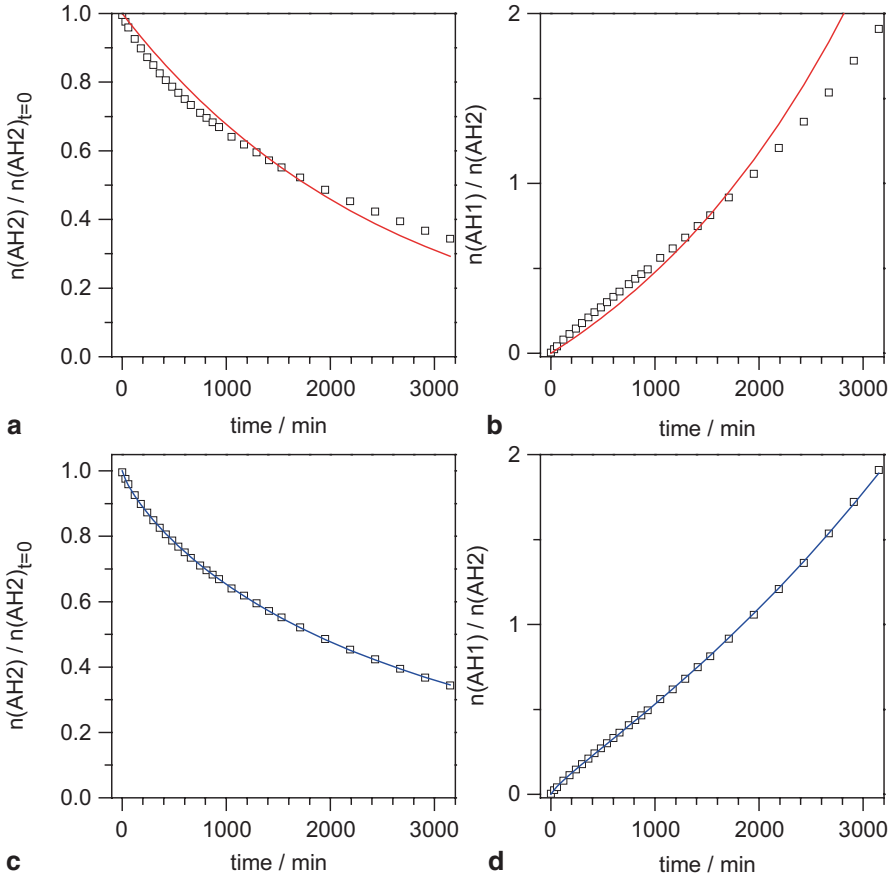


Fig. 7.19 Squares represent: (**a**, **c**) the evolution of the abundance of AH2; (**b**, **d**) the evolution of the population ratio of *AH1* and *AH2* rotamers with time of keeping the matrix in the *dark* at 13 K. Solid lines represent the best fits: (**a**, **b**) using the classical kinetics formulas and (**c**, **d**) using dispersive kinetics formulas [82]. (Reproduced with permission from [82], copyright © 2012 American Institute of Physics)

In a second experiment [70], narrowband tunable UV-induced mutual conformational interconversions between the two imino–oxo forms were observed, together with the oxo-hydroxy and amino-imino phototautomeric reactions. Irradiation of an argon matrix containing cytosine monomers with UV (λ=313 nm) laser light resulted in *syn*↔*anti* photoisomerizations between the two imino–oxo forms, whereas the substantially more populated amino–hydroxy and amino–oxo forms stayed intact (Fig. 7.20). Subsequent irradiation with shorter-wavelength UV (λ=311 nm) laser light led to two concomitant phototautomeric processes consuming the amino–oxo isomer (Fig. 7.21): (i) an oxo→hydroxy hydrogen-atom transfer photoprocess converting the amino–oxo form into the amino–hydroxy tautomer; (ii) an amino→imino hydrogen-atom transfer converting the amino–oxo form into

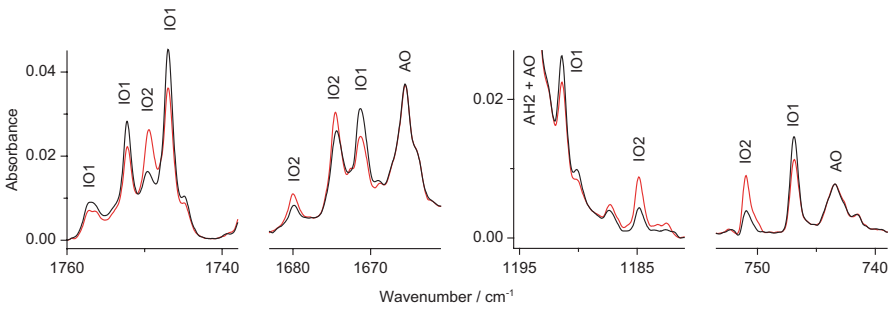


Fig. 7.20 Fragments of the infrared spectrum of cytosine monomers isolated in an Ar matrix: (*black*) recorded after deposition of the matrix; (*red*) recorded after irradiation of the matrix with UV (λ=312 nm) monochromatic laser light, showing the *syn*↔*anti* isomerization within the imino-oxo forms of cytosine. (Reproduced from [70], with permission from the PCCP Owner Societies)

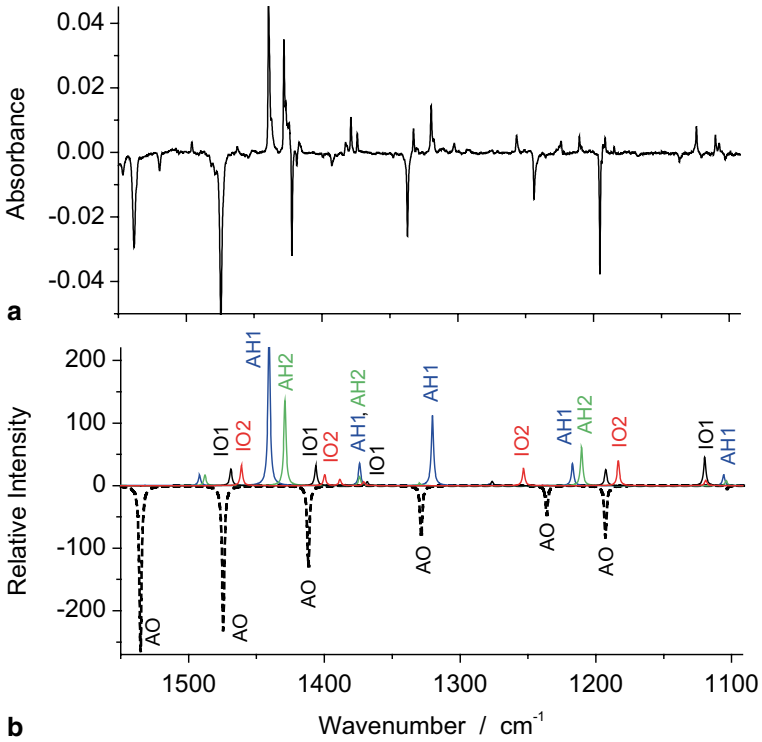


Fig. 7.21 Phototransformation of the AO tautomer of matrix-isolated cytosine into the *AH1*, *AH2*, *IO1* and *IO2* forms. (**a**) subtraction result: the experimental spectrum recorded after UV irradiation at 311 nm *minus* the spectrum recorded after the previous UV irradiation at 312 nm; (**b**) the spectra theoretically simulated (scaled) at the DFT(B3LYP)/6-31++G(d, p) level for *AH1* (*blue*), *AH2* (*green*), *IO1* (*black solid line*), *IO2* (*red*) and *AO* (*black dashed line*) isomers of cytosine. (Reproduced from [70], with permission from the PCCP Owner Societies)

the imino–oxo isomers. On the whole, all the five lower energy forms of cytosine could be experimentally observed.

7.5 Phenolic Compounds

Phenolic compounds have unique properties as antioxidants and free radical scavengers. Plant phenols, in particular, have attracted the attention of scientists interested in the elucidation of many aspects of their biological activity [83, 84]. The phenol moiety is also a fundamental building block of many other aromatic biomolecules, such as the aromatic amino acids tyrosine and phenylalanine and neurotransmitters (dopamine, among others). The parent phenol itself has long been receiving biomedical applications as disinfectant [85], and it is a precursor in the synthesis of aspirin. Nowadays, phenol derivatives like triclosan (or 5-chloro-2-(2,4-dichlorophenoxy) phenol) constitute major broad-spectrum antibacterial and antifungal agents found in a wide variety of consumer goods, like detergents, dish-washing liquids, soaps, deodorants, cosmetics, lotions, anti-microbial creams and toothpastes [86–90].

Phenols can also be extremely toxic. Approximately 1 g of parent phenol is enough to cause human death. During the Second World War, phenol injections were given to thousands of people in concentration camps by the Nazis (especially at Auschwitz-Birkenau), in what constitutes one of the darkest periods of the History of mankind.

Structurally, phenol is a simple molecule existing in a unique stable geometry in the electronic ground state, where the OH group stays in the plane of the aromatic ring. Substitutions in the ring may introduce asymmetry and lead to complex conformational landscapes. Isoeugenol (2-methoxy-4-(prop-1-enyl)phenol), for example, which is one of the most important plant phenols, has two geometric isomers, *E* and *Z*, differing in the relative position of the substituents around the C=C double bond of the propenyl moiety (Fig. 7.22). For each of these species, there are several conformers resulting from internal rotation about the single bonds connecting the $–CH=CHCH_3$, –OH and $–OCH_3$ groups to the aromatic ring. According to theoretical predictions, the *E*-type species with the intramolecular O−H···O hydrogen bond and with the *anti* arrangement of the propenyl substituent (E1 and E2 in Fig. 7.22) are the most stable forms of the isoeugenol molecule [91].

Upon isolation of isoeugenol in a cryomatrix, the four E1, E2, Z1 and Z2 forms of the compound were observed experimentally [91]. Due to the low energy barrier between the two *Z* forms (~2.8 kJ mol^{-1}) annealing of the matrices at 28 K led to observation of the Z1 → Z2 conversion. On the other hand, the E2 → E1 conformational conversion could not be observed experimentally by matrix annealing, since the energy barrier associated with this conversion is relatively high (>10 kJ mol^{-1}), thus precluding the conformational cooling.

Very interestingly, irradiation of matrix-isolated isoeugenol (both in argon and xenon matrices) in the 310–308 nm range was found to selectively convert the E1 form into the Z1 species (Fig. 7.23). Further irradiation at λ=305 nm induced

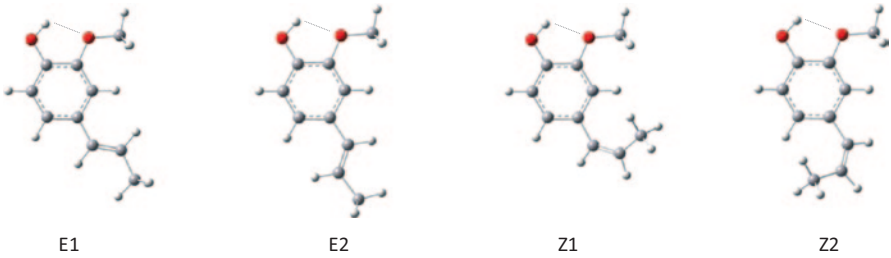


Fig. 7.22 Observed conformers of *E* and *Z* isomers of isoeugenol isolated in low temperature matrices

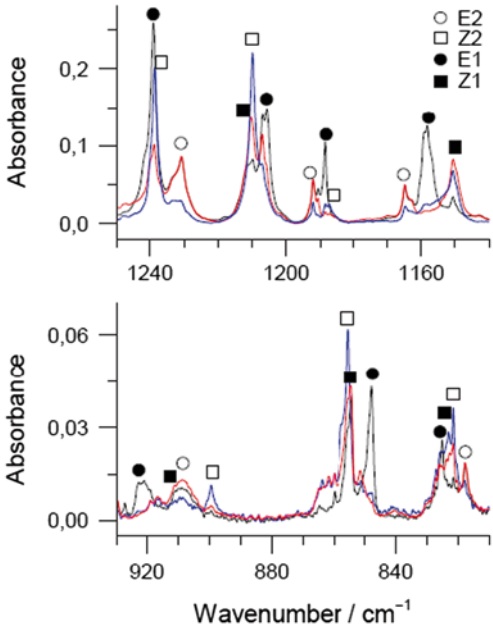


Fig. 7.23 Selected regions of the IR spectra of isoeugenol in argon matrix: freshly deposited matrix (*black traces*); after irradiation of a freshly deposited matrix at 310–308 nm (*red traces*) followed by irradiation at 302 nm (*blue traces*). (Adapted with permission from [91], copyright © 2012 American Chemical Society)

spectral changes that indicated occurrence in the matrix of the Z1 → E1 back reaction, thus providing an undoubted proof of the photoreversibility of the E1 ↔ Z1 transformation. On the other hand, when the sample was irradiated at $\lambda = 302$ nm, conversion of the E2 species into Z2 was observed (Fig. 7.23) [91]. These are interesting examples of conformer-selective photoisomerizations of a type similar to the triggering photochemical event in the vision process, involving conversion of 11-*cis*-retinal into all-*trans*-retinal by light [6].

When irradiation of matrix-isolated isoeugenol was carried out with UV light with $\lambda < 308$ nm, additional photoreactions were observed, leading to formation of two different long-chain ketene derivatives [91]. The primary photochemical process leading to these reactions is the H-atom detachment from the phenolic OH group, followed by reaction of the released hydrogen atom with the formed

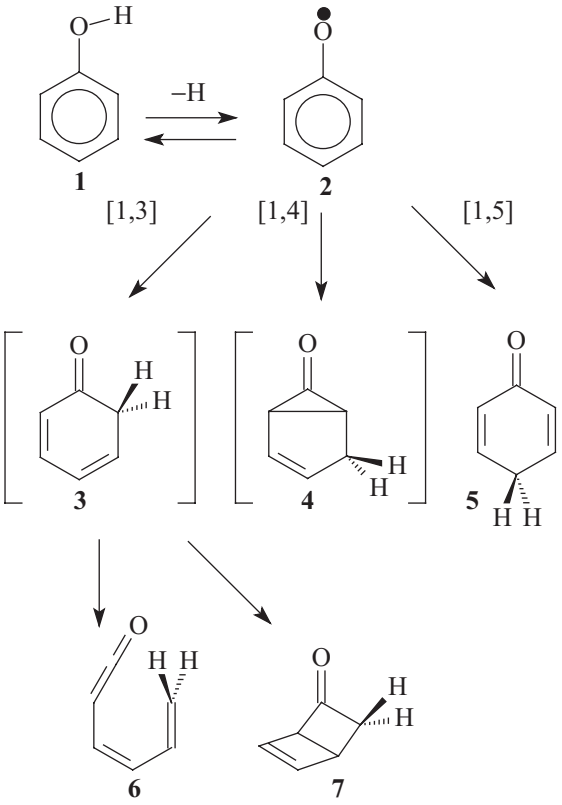


Fig. 7.24 Pathways of isomerizations induced by UV (λ = 275 nm) irradiation of matrix-isolated phenol. Structures in brackets were not observed. (Reproduced with permission from [92], copyright © 2012 American Institute of Physics)

phenolic radical. The two types of ketenes were formed depending on the position of the ring to where the hydrogen atom moves. Subsequent decarbonylation of these ketenes was also found to take place in the photolysed matrices. The same photochemical behavior was also observed for eugenol (2-metoxy-4-(prop-2-enyl) phenol) [91].

The H-atom detachment from the phenolic OH group seems in fact to be a quite general process for matrix-isolated phenols. For the parent phenol isolated in solid argon, UV (λ = 275 nm) excitation was found to lead to initial formation of the phenoxyl radical [92]. Following H-atom detachment from the phenolic OH group, the hydrogen atom may recombine in different ways with the phenoxyl radical, so that evidence of photogeneration of 2,5-cyclohexadienone and of an open-ring conjugated ketene and phenol Dewar isomers was found in the spectra recorded after UV irradiation of matrix-isolated phenol (Figs. 7.24 and 7.25) [92]. Upon shorter-wavelength UV excitation, further photoproducts such as cyclopentadiene, carbon monoxide, allene, acetylene, cyclobutadiene, and parent ketene were generated. No indication of photochemical channels leading to

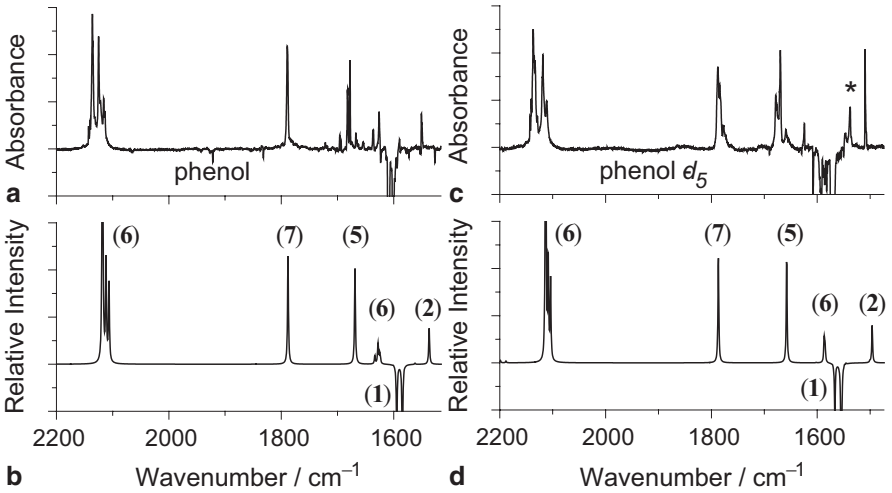


Fig. 7.25 Difference IR spectra showing bands of photoproducts generated upon UV λ = 275 nm irradiation of matrix-isolated **a** phenol and **c** phenol-d_5; **b**, **d** simulated difference spectra with calculated intensities weighted as [(**2**) × 50 + (**5**) × 15 + (**7**) × 15 + (**6**) × 20] − [(**1**) × 100], where [phenoxyl radical (**2**), 2,5−cyclohexadienone (*5*), open-ring ketene (*6*) and Dewar isomer (*7*), positive bands] grow at the expense of phenol (*1*), (negative bands, truncated). The open-ring ketene (*6*) was represented as a sum of its most stable four rotamers with equal weights (5 % each). The photoproduct band designated with *asterisk* in frame (**c**) was assigned to perdeuterated cyclobutadiene, formed in a secondary reaction. (Reproduced with permission from [92], copyright © 2012 American Institute of Physics)

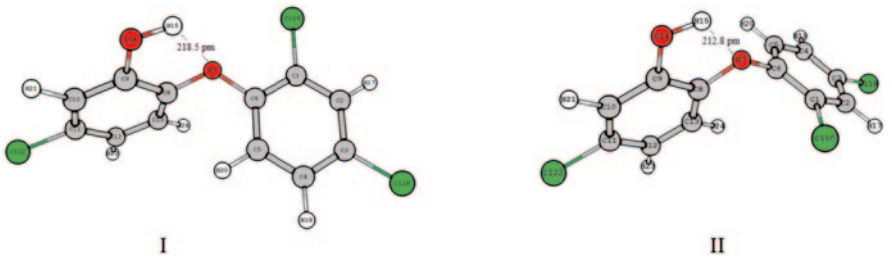


Fig. 7.26 Conformers of triclosan

a prefulvene isomer (4, in Fig. 7.24) was detected. The formation of the phenoxyl radical is particularly relevant, since this species is known to be involved in the phenol-induced toxicity [93].

In the case of triclosan, two types of photochemistries were observed, depending on the experimental conditions [94, 95]. Triclosan was shown to exist in two conformers (Fig. 7.26), which in the matrices were found to be present in a proportion similar to that predicted for the room temperature gas phase (I:II = 0.75). Upon broadband UV

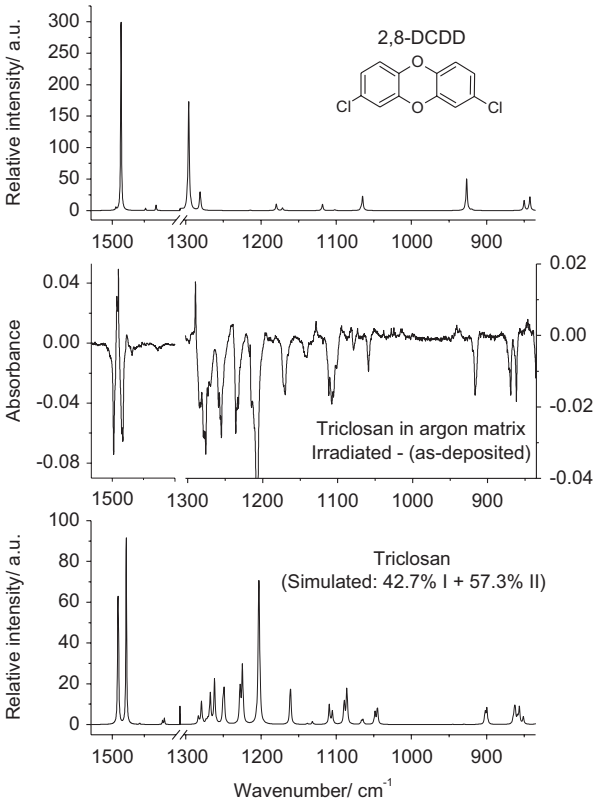


Fig. 7.27 Bottom to top: B3LYP/6-311++G(d, p) simulated IR spectrum of triclosan; difference spectrum: irradiated matrix by Xe arc lamp *minus* as-deposited matrix; simulated IR spectrum of 2,8-DCDD. (Reproduced from [94] with permission from Elsevier)

irradiation of the matrix-isolated triclosan with unfiltered light provided by a xenon arc-lamp, formation of the highly toxic 2,8-dichlorodibenzo-*p*-dioxin (2,8-DCDD) was observed (Fig. 7.27), together with HCl. This reaction seems to occur through initial photoproduction of the corresponding phenoxyl radical, thus closely following the photochemical behavior of parent phenol and isoeugenol [91, 92], and involves in the initial step participation of dissociative ($\pi\sigma^*$) excited states along the OH stretching coordinate, as observed previously for other phenol derivatives [96]. The photochemically detached hydrogen atom derived from triclosan may then react with the closest located chlorine atom in the triclosan molecule to yield HCl and a biradical species, which can subsequently undergo a ring-closure reaction by intramolecular recombination, leading to the observed 2,8-DCDD product. When irradiated with a narrowband laser-MOPO source ($\lambda = 290$ nm), the main photochemical product of triclosan is its open-ring ketene isomer. The ketene is first produced in conformations (type I forms) resembling more the reactant and, subsequently, converted to more stable, more extended conformations (type II forms), upon prolonged irradiation at $\lambda = 290$ nm (Fig. 7.28).

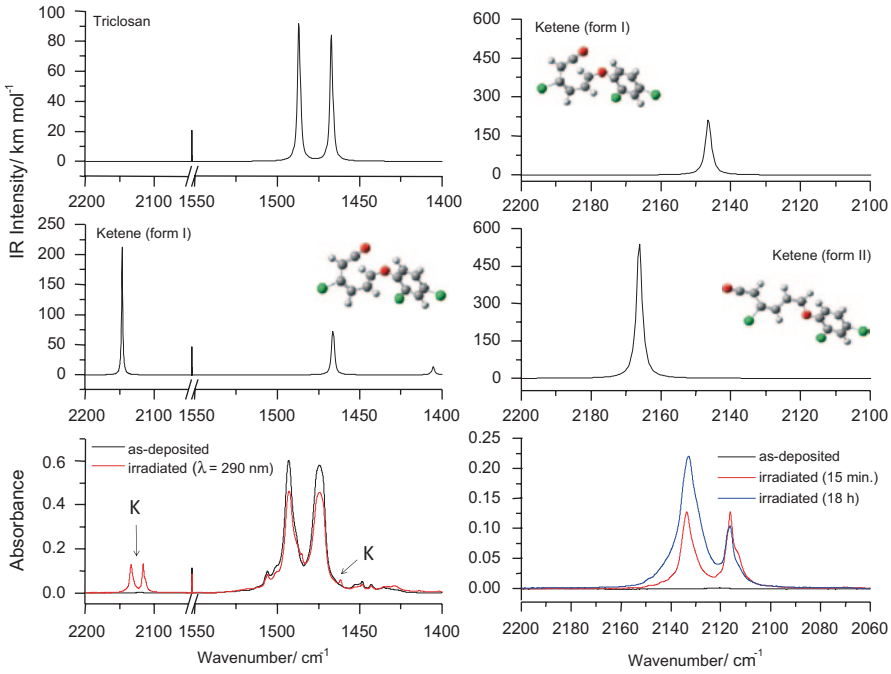


Fig. 7.28 B3LYP/6-311++G** simulated IR spectra of triclosan and its isomeric open-ring ketene (*forms type I and II*) and results of UV (λ = 290 nm) Laser-MOPO irradiation experiments in xenon matrix. *Left*: irradiated matrix for 15 min.; *K*, indicates ketene bands

7.6 Coumarins

Coumarin is the name given to the basic structural unit present in a large group of heterocyclic oxygen compounds that possess the benzopyran-2-one nucleus (*i.e.*, fused phenyl and α-pyrone rings; Fig. 7.29). They are of considerable biological and medical interest, showing important physiological effects (*e.g.*, anti-coagulant, anti-HIV, anti-tumor, anti-hypertension, anti-arrhythmia and anti-osteoporosis activities) and weak toxicity [97–101]. Natural products like esculetin, fraxetin, daphnetin and other related coumarin derivatives are also recognized to possess anti-inflammatory as well as antioxidant activities [102, 103].

Being structurally rigid, the coumarin moiety is not particularly challenging for structural studies. On the other hand, its photochemistry is rather versatile and has been the subject of many studies, in particular because these compounds are very efficient light emitters [104–107].

The photochemistry of coumarins is very much related with that of their heterocyclic ring fragment. As mentioned before, the α-pyrone ring exhibits a very interesting photochemistry. Unsubstituted α-pyrone was indeed one of the first heterocyclic compounds isolated in low temperature inert matrices in the early 1970s of the twentieth century [12, 13, 108]. In those studies, the isomeric open

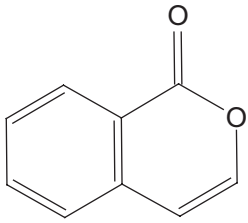


Fig. 7.29 Coumarin (or, benzopyran-2-one)

ring aldehyde-ketene was observed as the main product, being rapidly produced upon UV irradiation (λ = 313 nm) of the matrix. The IR bands assigned to the aldehyde-ketene were first interpreted on the basis of coexistence in the matrix of four isomers of the molecule, all exhibiting the *Z* orientation at the double C = C bond. The second photoproduct, the Dewar isomer of α-pyrone, was found to be generated very slowly under the used experimental conditions, a significant amount of the Dewar isomer being only observed after prolonged irradiation. Upon more energetic UV irradiation (λ > ~ 200 nm), the matrix-isolated Dewar isomer was found to decompose, with evolution of CO_2 and creation of anti-aromatic cyclobutadiene [12, 13, 108] (see Fig. 7.1).

More recently, the photochemistry of α-pyrone isolated in low temperature inert matrices was reinvestigated in our laboratory [14]. As expected, both the ring opening reaction leading to conjugated ketene photoproducts and valence isomerization to the Dewar form were observed, with the first type photochemistry dominating. However, the new experimental studies received support from extensive high-level theoretical calculations, not available at the time of the original studies on this compound. This advantage permitted to reveal some new facets in the photochemistry of α-pyrone. For example, photoproduction of the Z as well as E forms of the conjugated ketene was spectroscopically proven for the first time.

Several substituted α-pyrones, including some coumarins, were subsequentely studied [109–112]. It was shown that the Dewar isomer formation is much more effective in the case of the 4,6-dimethyl-α-pyrone, while the α-cleavage photoreaction, leading to the aldehyde-ketene, proceeds easier for unsubstituted α-pyrone [14, 109]. Upon UV-irradiation (λ > 200 nm) of coumarin isolated in solid argon, three main photoreactions were observed (Fig. 7.30): (a) decarboxylation of the compound and formation of benzocyclobutadiene and CO_2, with Dewar coumarin as intermediate; (b) ring-opening to its isomeric conjugated ketene; and (c) decarbonylation, leading to formation of a benzofuran/CO complex. Further decomposition of benzofuran to produce ethynol was also suggested [110]. On the other hand, in 3-acetamido-coumarin [111] only the ring-opening isomerization to the open ring ketene and the decarbonylation reactions were observed, while no photochemical production of the Dewar isomer occurs. This last result, follows the trend observed for 2-pyrone-3-carboxylate [112] and seems to be a quite general rule for matrix-isolated α-pyrones and coumarins bearing relatively volumous substituents at the position 3 of the pyrone ring, as a consequence of the unfavorable relaxation of the matrix around the guest molecule that would be required to accommodate the

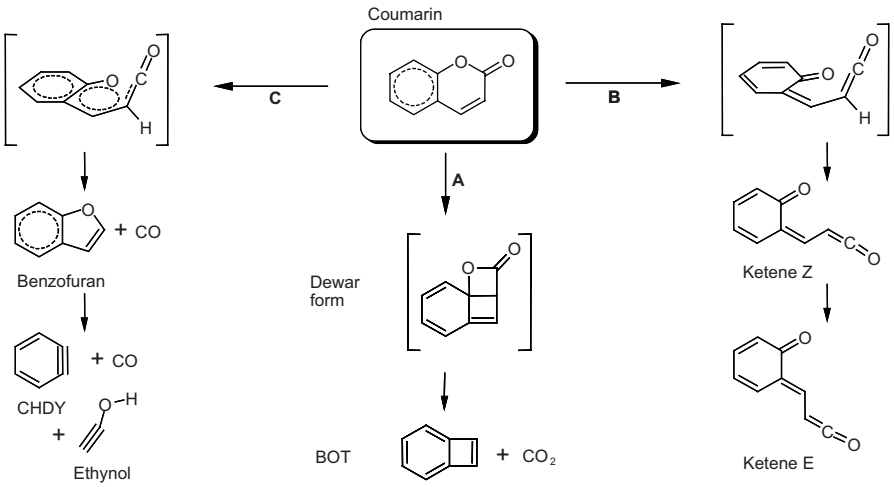


Fig. 7.30 Photolytic reaction pathways resulting from UV irradiation of matrix isolated coumarin. Structures in brackets are either not detected or transition states. *BOT* benzocyclobutadiene; *CHDY* cyclohexa-1,3-dien-5-yne (benzyne). (Reproduced from [110] with permission from John Wiley and Sons)

Dewar isomers of these compounds, whose structure deviates strongly from planarity, thus mismatching the primarily occupied matrix sites.

The main difference between the photochemistry shown by coumarin and α-pyrone is that the ring-opening reaction in coumarin occurs only upon excitation with $\lambda > 200$ nm, while in α-pyrone it starts readily upon irradiation with $\lambda > 337$ nm. This difference cannot be explained in terms of different energy gaps between the electronic ground states and the lowest excited states in two compounds. Indeed, in coumarin the two lowest singlet states were found to have energies of 29,100 (n,π*) cm^{-1} and 32,800 (π,π*) cm^{-1} [113, 114], while in α-pyrone the corresponding energies were comparable or slightly higher: 30,200 (n,π*) cm^{-1} and 35,000 (π,π*) cm^{-1} [115]. On the other hand, since to initiate the photochemical reaction an energy in the range $235 > \lambda > 200$ nm was found to be necessary, it can be concluded that the target electronic excitation corresponds to the band at *ca.* 210 nm (~47 600 cm^{-1}) observed in the UV absorption spectrum of coumarin (in ethanol) at room temperature [116]. According to calculations in Reference [113], this band relates to S_5, that is, four singlet states and five triplet states are situated below this level. The following main factors can justify the more difficult α-cleavage in coumarin when compared to α-pyrone: (i) In coumarin, the presence of the additional phenyl ring results in an increased density of states in the vicinity of S_1 and S_2. In this molecule, S_1 (n, π*) and T_3 (n, π*) are almost isoenergetic: $29{,}100 \pm 500$ cm^{-1} and $28{,}000 \pm 300$ cm^{-1}, respectively [113]. This can result in a very effective intersystem crossing and subsequent quenching, *via* T_2, to T_1 ($22{,}000 \pm 300$ cm^{-1}). The lowest triplet state is probably too low in energy to induce the ring cleavage. (ii) The larger number of atoms in coumarin results in the increase of the vibrational degrees

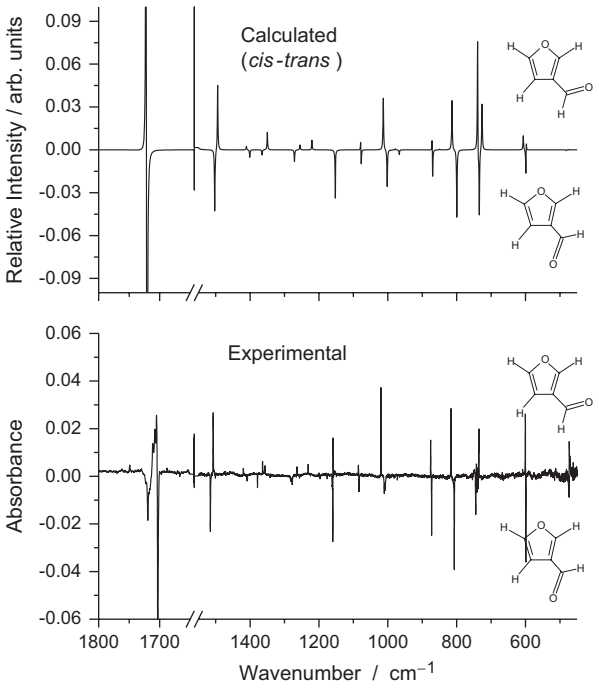


Fig. 7.31 *Bottom:* difference infrared spectrum (irradiated *minus* as-deposited matrix) of 3FA, showing the *trans*→*cis* photoisomerization. *Top:* B3LYP/6 311++G(d, p) simulated *cis*-*trans* difference spectrum. (Adapted with permission from [117], copyright © 2010 American Chemical Society)

of freedom that can serve as intermediate states for the intramolecular processes of non-radiative dissipation of energy. (iii) In α-pyrone, the ring-opening reaction results in the increase of the number of double bonds, which is a stabilizing factor in terms of the electronic π-system. On the other hand, in coumarin the opening of the pyrone ring reduces the aromaticity of the neighboring phenyl ring (the relative energies of the open-ring molecules relatively to the closed-ring counterparts are equal to *ca.* 75 kJ mol^{-1} in α-pyrone and *ca.* 140 kJ mol^{-1} in coumarin [110]). (iv) Competitive photolysis processes are more important in coumarin. Indeed, the present experimental data indicate that, along with the ring-opening reaction, the decarboxylation (through the Dewar isomerization) and decarbonylation reactions do also occur. In α-pyrone, because the ring-opening reaction occurs at lower excitation energies than the isomerization to the Dewar form, these competing mechanisms are not operative (or not efficient enough).

Conformational photoisomerization of the 5-membered-ring compound analogous of α-pyrone, 3-furaldehyde, was found to be particularly interesting [117]. 3-Furaldehyde has two conformers differing in the orientation of the aldehyde group (Fig. 7.31), the *trans* form being *ca.* 4 kJ mol^{-1} more stable than the *cis* form. Accordingly, the two conformers could be trapped in an argon matrix from the compound room temperature gas phase in proportion ~7:1. UV-irradiation ($\lambda > 234$ nm) of the matrix-isolated compound resulted in partial *trans*→*cis* isomerization (Fig. 7.31), which ended at a photostationary state with the *trans*/*cis* ratio being *ca.* 2:1.

The obtained population ratio at the photostationary state was explained based on results of time-dependent DFT calculations. The potential energy profiles of the excited states calculated as a function of the aldehyde group rotation showed that all of S_1, S_2, and S_3 states have a minimum at the *cis* conformation, which is, in those states, more stable than the *trans* conformation. In addition, the bright S_2 state shows a third minimum (global minimum) at a nearly perpendicular conformation of the aldehyde relatively to the molecule ring (~97°). Excitation of both conformers with $\lambda > 234$ nm provides the excess energy of around 5.3 eV, which is well above all the barriers on the S_2 surface and can be followed by fast internal rotation in the S_2 excited state to the ~97° minimum. Relaxation of S_2 from the ~97° minimum can then produce either *cis* or *trans* ground state conformers by internal rotation in S_0. However, since the torsional coordinate at the minimum in S_2 equals ~97º, it is shifted in the direction of the *trans* conformer, thus leading to the observed more favourable relaxation to this form [117].

Acknowledgements The authors thank all their colleagues involved in the studies described in more detail in this Chapter. Finantial support from *Fundação para a Ciência e a Tecnologia* (FCT – Portugal) is acknowledged.

References

1. Rosemeyer H (2004) Chem Biodiversity 1:361
2. Fischer E (1899) Ber Deutsch Chem Gesellschaft 32:2550
3. Yamada H, Okamoto T (1972) Chem Pharm Bull 20:623
4. Movassaghi M, Hill MD (2006) J Am Chem Soc 128:14254
5. Nelson DL, Cox MM (2008) Principles of biochemistry, 5th edn. W. H. Freeman and Company, pp. 272–274
6. Schertler GFX (2008) Nature 453:292
7. Saunders J (2003) Nature Rev Microbiol 1:6
8. Sabbert D, Engelbrecht S, Junge W (1996) Nature 381:623
9. Whittle E, Dows DA, Pimentel GC (1954) J Chem Phys 22:1943
10. Norman I, Porter G (1954) Nature 174:508
11. Lin CY, Krantz A (1972) J Chem Soc Chem Commun 1111/1112:1316
12. Chapman OL, McIntosh CL, Pacansky J (1973) J Am Chem Soc 95:244
13. Pong RGS, Shirk JS (1973) J Am Chem Soc (95) 248
14. Breda S, Reva I, Lapinski L, Fausto R (2004) Phys Chem Chem Phys 6:929
15. Rochkind MM (1967) Anal Chem 39:567
16. Perutz RN, Turner JJ (1973) J Chem Soc Faraday Trans 2(69):452
17. Barnes AJ, Bignall JC, Purnell CJ (1975) J Raman Spectrosc 4:159
18. Fausto R, Gómez-Zavaglia A (2009) Light-induced reactions in cryomatrices. In: Abini A (ed) Specialist periodical reports—photochemistry, vol 37. The Royal Society of Chemistry Publishing, pp 72–109
19. Fausto R, Gómez-Zavaglia A (2010) Light-induced reactions in cryomatrices. In: Abini A (ed) Specialist periodical reports—photochemistry, vol 38. The Royal Society of Chemistry Publishing, pp 37–66
20. Fausto R, Gómez-Zavaglia A (2011) Light-induced reactions in cryomatrices. In: Abini A (ed) Specialist periodical reports—photochemistry, vol 39. The Royal Society of Chemistry Publishing, pp 1–29

21. Gómez-Zavaglia A, Fausto R (2003) Phys Chem Chem Phys 5:3154
22. Stepanian SG, Reva ID, Radchenko ED, Rosado MTS, Duarte MLTS, Fausto R, Adamowicz L (1998) J Phys Chem A 102:1041
23. Gómez-Zavaglia A, Fausto R (2003) Vibrat Spectrosc 33:105
24. Gómez-Zavaglia A, Reva ID, Fausto R (2003) Phys Chem Chem Phys 5:41
25. Rosado MTS, Duarte MLTS, Fausto R (1998) Vibrat Spectrosc 16:35
26. Suenram RD, Lovas FJ (1978) J Mol Spectrosc 72:372
27. Suenram RD, Lovas FJ (1980) J Am Chem Soc 102:7180
28. Schäfer L, Sellers HL, Lovas FJ, Suenram RD (1980) J Am Chem Soc 102:6566
29. Godfrey PD, Brown RD (1995) J Am Chem Soc 117:2019
30. Lovas FJ, Kawashima Y, Grabow J-U Suenram RD, Freser GT, Hirota E (1995) AstrophysJ 455:201
31. Iijima K, Tanaka K, Onuma S (1991) J Mol Struct 246:257
32. Borba A, Gómez-Zavaglia A, Fausto R (2006) J Mol Struct 794:196
33. Kaczor A, Reva ID, Proniewicz LM, Fausto R (2007) J Phys Chem A 111:2957
34. Stepanian SG, Reva ID, Radchenko ED, Adamowicz L (1998) J Phys Chem A 102:4623
35. Stepanian SG, Reva ID, Radchenko ED, Adamowicz L (2001) J Phys Chem A 105:10664
36. Jarmelo S, Lapinski L, Nowak MJ, Carey PR, Fausto R (2005) J Phys Chem A 109:5689
37. Jarmelo S, Fausto R (2006) J Mol Struct 786:175
38. Kaczor A, Reva ID, Proniewicz LM, Fausto R (2006) J Phys Chem A 110:2360
39. Petrella RJ, Karplus M (2004) Proteins: Struct Funct Bioinf 54:716
40. Stewart RF (1967) Acta Cryst 23:1102
41. Szczesniak M, Nowak MJ, Rostkowska H, Szczepaniak K, Person WB, Shugar D (1983) J Am Chem Soc 105:5969
42. Barnes AJ, Stuckey MA, Le Gall L (1984) Spectrochim Acta A 40:419
43. Nowak MJ (1984) J Mol Struct 116:49
44. Ivanov AY, Plokhotnichenko AM, Radchenko ED, Sheina GG, Blagoi YP (1995) J Mol Struct 372:91
45. Les A, Adamowicz L, Nowak MJ, Lapinski L (1992) Spectrochim Acta A 48:1385
46. Graindourze M, Smets J, Zeegers-Huyskens T, Maes G (1990) J Mol Struct 222:345
47. Szczepaniak K, Szczesniak M, Nowak MJ, Scott I, Chin S, Person WB (1984) Int J Quant Chem 18:547
48. Feyer V, Plekan O, Richter R, Coreno M, de Simone M, Prince KC, Trofimov AB, Zaytseva IL, Schirmer J (2010) J Phys Chem A 114:10270
49. Peeters Z, Botta O, Charnley SB, Ruiterkamp R, Ehrenfreund P (2003) Astrophys J 593:L129
50. Perun S, Sobolewski AL, Domcke W (2006) J Phys Chem A 110:9031
51. Schultz T, Samoylova E, Radloff W, Hertel IV, Sobolewski AL, Domcke W (2004) Science 306:1765
52. Martins Z, Botta O, Fogel ML, Sephton MA, Glavin DP, Watson JS, Dworkin JP, Schwartz AW, Ehrenfreund P (2008) Earth Planet Sci Lett 270:130
53. Stokes PG, Schwartz AW (1979) Nature 282:709
54. Stokes PG, Schwartz AW (1981) Geochim Cosmochim Acta 45:563
55. van der Velden W, Schwartz AW (1977) Geochim Cosmochim Acta 41:961
56. Radchenko ED, Plokhotnichenko AM, Sheina GG, Blagoi YP (1983) Biofizika 28:923
57. Nowak MJ (1989) J Mol Struct 193:35
58. Radchenko ED, Plokhotnichenko AM, Sheina GG, Blagoi YP (1984) Biofizika 29:553
59. Stepanian SG, Sheina GG, Radchenko ED, Blagoi YP (1985) J Mol Struct 131:333
60. Sheina GG, Radchenko ED, Stepanian SG, Blagoi YP (1986) Stud Biophys 114:123
61. Nowak MJ, Lapinski L, Kwiatkowski JS (1989) Chem Phys Lett 157:14
62. Nowak MJ, Lapinski L, Kwiatkowski JS, Leszczynski J (1996) J Phys Chem A 100:3527
63. Szczesniak M, Kwiatkowski JS, Kubulat K, Szczepaniak K, Person WB (1988) J Am Chem Soc 110:8319
64. Nowak MJ, Lapinski L, Fulara J (1989) Spectrochim Acta A 45:229

65. Szczepaniak K, Szczesniak M (1987) J Mol Struct 156:29.66. Kwiatkowski JA, Pullman B (1975) Adv Heterocycl Chem 18?:199
67. Fazakerley GV, Gdaniec Z, Sowers LC (1993) J Mol Biol 230:6
68. Khuu P, Ho PS (2009) Biochemistry 48:7824
69. Griffiths AJF, Miller JH, Suzuki DT, Lewontin RC, Gelbart WM (2000) An introduction to genetic analysis, 7th edn. W. H. Freeman, New York (U.S.A.)
70. Lapinski L, Reva I, Nowak MJ, Fausto R (2011) Phys Chem Chem Phys 13:9676
71. Fogarasi G (2002) J Phys Chem A 106:1381
72. Piacenza M, Grimme S (2004) J Comput Chem 25:83
73. Wolken JK, Yao C, Turecek F, Polce MJ, Wesdemiotis C (2007) Int J Mass Spectrom 267:30
74. Trygubenko SA, Bogdan TV, Rueda M, Orozco M, Luque FJ, Sponer J, Slavicek P, Hobza P (2002) Phys Chem Chem Phys 4:4192
75. Kobayashi R (1998) J Phys Chem 102:10813
76. Brown RD, Godfrey PD, McNaughton D, Pierlot AP (1989) J Am Chem Soc 111:2308
77. Feyer V, Plekan O, Richter R, Coreno M, Vall-Llosera G, Prince KC, Trofimov AB, Zaytseva IL, Moskovskaya TE, Gromov EV, Schirmer J (2009) J Phys Chem A 113:5736
78. Lapinski L, Nowak MJ, Reva I, Rostkowska H, Fausto R (2010) Phys Chem Chem Phys 12:9615
79. Maçôas EMS, Kriachtchev L, Pettersson M, Fausto R, Räsänen M (2003) J Am Chem Soc 125:16188
80. Maçôas EMS, Fausto R, Pettersson M, Khriachtchev L, Räsänen M (2000) J Phys Chem A 104:6956
81. Fausto R, Maçôas EMS (2001) J Mol Struct 563/564:29
82. Reva I, Nowak MJ, Lapinski L, Fausto R (2012) J Chem Phys 136:064511
83. Southwell IA, Russell MF, Davies NW (2011) Flav Fragr J 26:336
84. Ayala-Zavala JF, González-Aguilar GA, Del-Toro-Sanchez L (2009) J Food Sci 74:84
85. Cook AM (1960) J Pharm Pharmacol 12:19T
86. Council Directive 76/768/EEC on the Approximation of the Laws of the Member States Relating to Cosmetic Products, European Commission, Brussels, Belgium, 1999
87. Furia TE, Schenkel AG (1968) Soap Chem Special 44:47
88. Bhargava HN, Leonard PA (1996) Am J Infect Control 24:209
89. McMurry LM, Oethinger M, Levy SB (1998) Nature 394:531
90. Iconomopoulou SM, Andreopoulou AK, Soto A, Kallitsis JK, Voyiatzis GA (2005) J Contr Release 102:223
91. Krupa J, Olbert-Majkut A, Reva I, Fausto R Wierzejewska M (2012) J Phys Chem B DOI: 10.1021/jp306339 g (in the press)
92. Giuliano BM, Reva I, Lapinski L, Fausto R (2012) J Chem Phys 136:024505
93. Hanscha C, McKarnsb SC, Smith CJ, Doolittle DJ (2000) Chem-Biol Interact 127:61
94. Kuş N, Reva I, Bayari S, Fausto R (2012) J Mol Struct 1007:88
95. Kuş N, Bayari S, Reva I, Fausto R (2012) Proceedings of the XXIV IUPAC Symposium on Photochemistry, Coimbra, 15–20 July 2012. p 95
96. Ashfold MNR, Devine AL, Dixon RN, King GA, Nix MGD, Oliver TAA (2008) Proc Natl Acad Sci USA 105:12701
97. Bultink IEM, Lems WF, Kostense PJ, Dijkmans BAC, Voskuyl AE (2005) Arthritis Rheum 52:2044
98. O'Reilly RA (1967) J Clin Invest 46:829
99. Zhao H, Neamati N, Hong HX, Mazumder A, Wang SM, Sunder S, Milne GWA, Pommier Y, Burke TR (1997) J Med Chem 40:242
100. Hirsh J, Dalen JE, Anderson DR, Poller L, Bussey H, Ansell J, Deykin D (2001) Chest 119:8
101. Kashman Y, Gustafson KR, Fuller RW, Cardellina JH, McMahon JB, Currens MJ, Buckheit RW, Hughes SH, Cragg GM, Boyd MR (1992) J Med Chem 35:2735

102. Fylaktakidou KC, Hadjipavlou-Litina DJ, Litinas KE, Nicolaides DN (2004) Curr Pharm Des 10:3813
103. Gacche RN, Gond DS, Dhole NA, Dawane BS (2006) J Enzyme Inhib Med Chem 21:157
104. Maeda M (1984) Laser dyes. Academic Press, New York (U.S.A.)
105. Kumar S, Giri R, Mishra SC, Machwe MK (1995) Spectrochim Acta A 51:1459
106. Raikar US, Renuka CG, Nadaf YF, Mulimani BG (2006) J Mol Struct 787:127
107. Christie RM (2001) Colour chemistry. Royal Society of Chemistry, Cambridge (UK)
108. Chapman OL, McIntosh CL, Pacansky J (1973) J Am Chem Soc 95:614
109. Breda S, Lapinski L, Reva I, Fausto R (2004) J Photochem Photobiol A 162:139
110. Kuş N, Breda S, Reva I, Tasal E, Ogretir C, Fausto R (2007) Photochem Photobiol 83:1237 (Erratum: ibidim, 83, 1541)
111. Kuş N, Breda S, Fausto R (2009) J Mol Struct 81:924–926
112. Reva ID, Nowak MJ, Lapinski L, Fausto R (2008) Chem Phys Lett 452:20
113. Seixas de Melo JS, Becker RS, Maçanita AL (1994) J Phys Chem 98:6054
114. Fausto R, Breda, S, Kuş N (2008) J Phys Org Chem 21:644
115. Seixas de Melo JS, Quinteiro G, Pina J, Breda S, Fausto R (2001) J Mol Struct 565/566:59
116. Song PS, ordon III WH (1970) J Phys Chem 74:4234
117. Kuş N, Reva I, Fausto R (2010) J Phys Chem A 114:12427

Chapter 8
Two-Dimensional Mid-Infrared Correlation Spectroscopy in Protein Research

Bogusława Czarnik-Matusewicz and Young Mee Jung

Abstract Representative results showing the position and strength of two-dimensional (2D) correlation spectroscopy in protein research are surveyed in this article. Special emphasis is placed on infrared spectroscopy. Different types of external perturbations that are particularly useful for exploring properties of proteins are discussed. Most promising developments in 2D correlation spectroscopy are demonstrated through results obtained for protein systems. The aim of this article has been to present 2D correlation spectroscopy as a simple method that significantly improves information about protein structure gained from infrared spectra.

Keywords 2D correlation spectroscopy • Infrared spectroscopy • Protein structure • Amide I • Heterospectral correlation

8.1 Introduction

Combined with "protein research," the following three phrases could confuse a reader who is unfamiliar with the subjects and the terminology used: "Two-dimensional infrared (2D IR) spectroscopy", "Correlation-2D IR spectroscopy" and "Two-dimensional correlation spectroscopy (2DCoS)". Additionally, the apparent similarities in the graphical presentations of the results from each of these spectroscopies could be very misleading. Generally, all of the phrases relate to methods that have been inspired by the two-dimensional techniques first developed in the field of nuclear magnetic resonance. However, 2D IR spectroscopies and 2D correlation spectroscopies are fundamentally different because they are based on nonlinear and linear interactions of electromagnetic waves with matter, respectively, as was underlined by Noda and Ozaki [1]. According to [2], the method of 2D IR spectroscopy concerns a nonlinear interaction of light and matter in which measurements

B. Czarnik-Matusewicz (✉)
Faculty of Chemistry, University of Wrocław, F. Joliot-Curie 14,
50–383, Wrocław, Poland
e-mail: boguslawa.czarnik@gmail.com

Y. M. Jung
Department of Chemistry, Kangwon National University,
Chuncheon, 200–701 Korea

M. Baranska (ed.), *Optical Spectroscopy and Computational Methods in Biology and Medicine,* Challenges and Advances in Computational Chemistry and Physics 14,
DOI 10.1007/978-94-007-7832-0_8, © Springer Science+Business Media Dordrecht 2014

are performed in two independent time dimensions to derive two-dimensional (2D) spectra in the frequency domain. The extent of the vibrational spectrum over the two frequency axes reports on how the excitation of a vibration with a given frequency influences all the other vibrations within a detection window after a waiting period. Such a spectrum gives insight into the couplings between the different excitations of the system under study and the time evolution of these couplings. Therefore, the 2D IR measurement must be performed within a picosecond or faster, i.e., on a time scale that is fast compared to most protein dynamics. If the waiting time is varied, the correlation between the excitation and detection frequencies is studied and is the subject of correlation-2D IR spectroscopy. In contrast to the phrases "2D IR spectroscopy" and "correlation-2D IR spectroscopy," which refer to nonlinear optical 2D spectroscopic techniques, the phrase "2D correlation spectroscopy (2DCoS)" refers to a simple mathematical procedure used to generate 2D correlation maps from spectra that are collected during the interaction of different electromagnetic probes with an externally perturbed system composed of absorbing molecules.

8.2 Protein Research and 2DCoS

In 1985, Yang et al. [3] have shown that when processed by Fourier self-deconvolution, the resolution-enhanced infrared spectra of a protein substantially improved quantitative estimates of the proportion of each secondary conformation in protein, which was measured both as a powder and in an aqueous solution. These results initiated a "golden period" for infrared spectroscopy in protein research; this period has been punctuated by results showing that the application of resolution enhancement procedures such as Fourier self-deconvolution, curve fitting, and second derivative spectroscopy could provide varying degrees of success [4].

In parallel to studies of proteins by infrared spectroscopy, the two-dimensional correlation method has been developed. One year after Yang's paper [3], in 1986, Noda presented the basic concept for perturbation-based two-dimensional spectroscopy applicable to infrared data [5]. In 1993, to analyze a set of dynamic spectra obtained under an external perturbation, Noda published a generalized 2D correlation spectroscopy [6] that attracted the attention of researchers working in different subjects. As a model-free approach, 2DCoS has been an alternative to the model-based Fourier self-deconvolution, curve fitting, and second derivative methods. In an unbiased manner, 2DCoS allows the detection of the location of highly overlapped bands and the separation of these bands into components that could be assigned to individual absorbers, i.e., the α-helix, β-sheet, PPII, β-turn, unordered elements, and side chains. Because of synchronous and asynchronous correlation among the intensities of the sub-bands, these sub-bands could be very reliably assigned to different functional groups, structure elements, and components of the studied protein system. As the data subjected to 2DCoS are composed from a series of spectra of the protein system measured as a function of a perturbation, 2DCoS opens a new possibility for monitoring the evolution of spectral changes that arise from forms that are active

over a narrow range of perturbations. The main advantage of 2DCoS is that 2DCoS probes the specific sequence of spectral intensity changes that take place during the perturbation-controlled measurement. Whenever 2DCoS is applied to protein studies the sequence of the secondary structure changes on the perturbation axes is predicted from the synchronous and asynchronous changes according to Noda's rules [1, 6]. The sequence of events related to infrared spectra makes it possible to elucidate the nature of the structural transitions and makes an important contribution to finding the non-native local minima in which partially folded proteins can become trapped. The fact that infrared 2DCoS allows the evolution of protein structures to be explored has helped to maintain and has even intensified the interest of researchers in the use of infrared spectroscopy to solve an old problem in protein chemistry.

Recently, interest in vibrational spectroscopy has additionally been motivated by the fact that this technique belongs to limited group of methods employed in studies of amyloid fibril formation. Although the studies are strongly supported by molecular dynamics (MD) simulations, reliable results cannot yet be obtained because this method is still limited to a set of small, simple protein folds [7]. For proteins with more than approximately 100 residues, experiments generally reveal that one (or more) intermediate is significantly populated during the folding process [8]. The problem is that "We do not have a quantitative microscopic understanding of the folding routes or transition states" [9]. Without knowledge of the structure of the intermediates, the mechanisms through which the fibrillation processes proceeds cannot be determined. Moreover, the structural information for these forms is not always available or of high resolution from X-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy. These techniques are especially ill-suited for studying noncrystalline and insoluble full-length fibrils [10]. Optical spectroscopies supported by two-dimensional correlation spectroscopy are potential alternatives to X-ray and NMR methods in these studies [11]. In particular, infrared spectroscopy is very well suited to following the processes of fibril formation from β-type proteins [12]. Bands between 1,611–1,630 cm^{-1} are regarded as hallmarks of the β-conformation of amyloid fibrils. In 2005, Meier in the article titled "Vibrational spectroscopy: a 'vanishing' discipline?" answered "no" and argued that 2DCoS and infrared imaging have brought many new opportunities, so vibrational spectroscopy is far from 'dead' [13]. Currently, vibrational spectroscopy combined with 2DCoS represents a particularly powerful tool in the investigation of amyloid fibril formation, which is implicated in misfolding diseases such as type II diabetes, Alzheimer's, and Parkinson's, which makes vibrational spectroscopy more alive than ever.

8.3 Generalized 2DCoS

Since 1993, the generalized version of 2DCoS has found numerous applications in a variety of fields, as described by Noda and Ozaki in the "Handbook of Vibrational Spectroscopy" [14, 15], in the Frontiers of Molecular Spectroscopy [16] and in the text book "Two-dimensional Correlation Spectroscopy—Applications in Vibrational

and Optical Spectroscopy" [1], which is by far the most comprehensive resource on 2DCoS. Moreover, Noda published several review papers [17–21] that give the detailed state of the applications of 2DCoS after each 2DCoS symposium. 2DCoS has also been popularized among spectroscopists in Japan and China through a review published by Ozaki and co-workers in the Bulletin of the Chemical Society of Japan [22] and in Science in China [23]. Readers who are interested in computational details should refer to the above cited review works and especially to the textbook [1].

A brief introduction to generalized 2D correlation spectroscopy with computational details for pretreatment and normalization of raw data with PLS_Toolbox (Eigenvector Research, Inc.), and calculation of the synchronous and asynchronous spectra with command-line functions in MATLAB are included in Czarnecki's review [24]. He has provided a survey of applications of 2DCoS to studies of the molecular structure and hydrogen bonding properties of basic molecules and their binary mixtures with water. The MATLAB commands for calculations of 2D spectra are also presented in [25]; moreover, software for 2DCoS on the Ozaki Lab website (http://science.kwansei.ac.jp/~ozaki/) is available for download for free.

It is strongly recommended to write own program for 2DCoS from calculation of the synchronous and asynchronous spectra up to the final graphical presentation. Each program should be verified on simulated data that have well-established intensity changes and a priori known temporal order [26]. This procedure allows to compare the sequence of changes obtained from the 2D analysis with a well-known profile and helps new 2D users understand the so-called Noda's rules and avoid misinterpretation of 2DCoS results. During the last few years several papers describing artifacts associated with the visualization and interpretation of 2D correlation maps were published [27–29]. Some of these objections [27] may result from false interpretations of simulated data [30]. An essential contribution to solving the problem has come from Lednev's group; the latter have demonstrated how some of the 'artifacts' of 2DCoS could be transformed into strengths of this method [31–33].

The fundamental aspects of generalized 2DCoS are combined into one scheme presented in Fig. 8.1. This figure shows the basic experimental and computational stages of the two-dimensional correlation analysis, which can be easily performed whenever a set of spectral responses for a protein system has been collected as a function of a perturbation variable.

Many details concerning the application of generalized 2D correlation spectroscopy to different problems and the combination of 2DCoS with multivariate chemometric techniques, which are widely used for the analysis of spectral data, can be found in [34]. A detailed introduction to generalized two-dimensional correlation spectroscopy, especially in the context of analysis of the infrared reflection–absorbance spectroscopy (IRRAS) spectra of biologically relevant molecules at the air–water interface perturbed by surface pressure or temperature, has been presented by Dluhy et al. [35].

Since the first 2DCoS symposium in 1999, a rapid increase in the number of applications of 2DCoS in investigations of the intriguing properties of biomolecular systems measured by different techniques has been observed. Many of these applications concern studies of protein systems through the following optical spec-

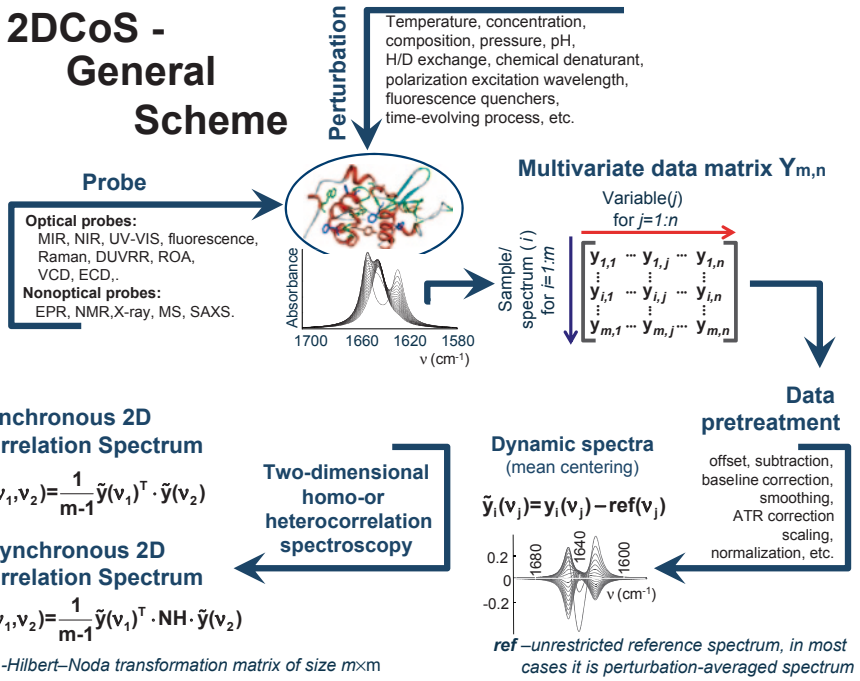


Fig. 8.1 General scheme to obtain 2D correlation spectra

troscopy methods: mid-IR [36–89], near-IR [36, 38, 90–99], vibrational circular dichroism (VCD) [43, 100, 101], electronic circular dichroism (ECD) [41, 102], Raman [40, 41, 103–107], Raman optical activity (ROA) [105, 107–111], deep UV resonance Raman (DUVRR) [28, 31, 32, 112–114], and fluorescence [115–118]. Most of these applications use mid-infrared spectroscopy, and this use explains the general popularity of this method in the investigation of different quantitative problems in biomolecular systems.

A few reviews devoted to the application of Raman, near- and mid-IR 2DCoS to protein research were given in [32, 119–121] along with detailed descriptions of the background and of the computational procedures. Additionally, in chapter 12 in the text book [1], the authors focus on infrared 2DCoS as a powerful tool in the examination of three selected protein systems. The most recent review of the application of vibrational 2DCoS to biological compounds was provided by Ozaki and Šašic [122] where results from [52] were presented in detail.

8.3.1 Pretreatment stage of 2DCoS analysis

Measurements as a function of the different perturbations listed in Fig. 8.1 do not require advanced equipment and can be conducted on an ordinary commercial IR

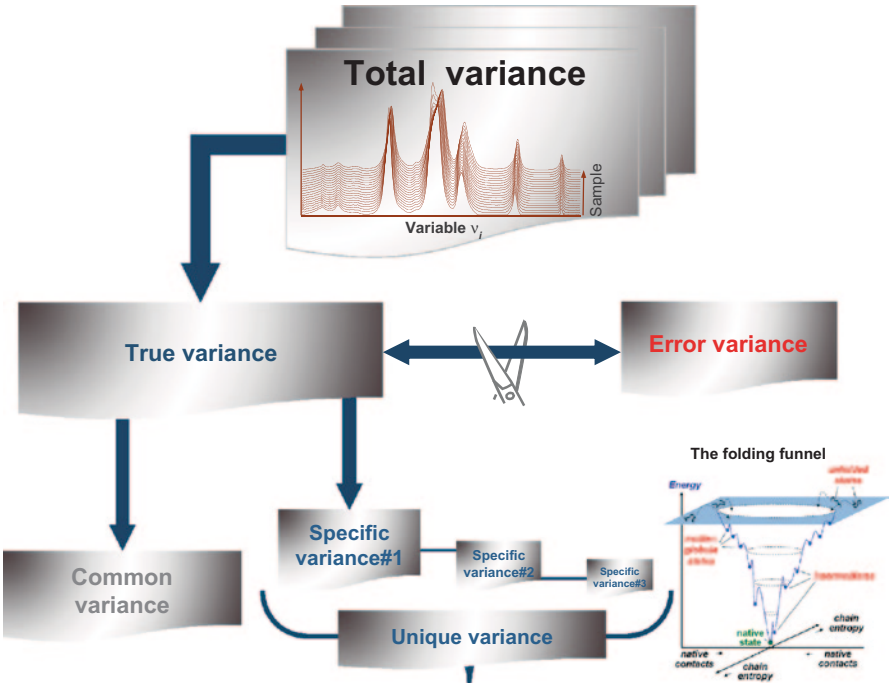


Fig. 8.2 Scheme for the separation of the total variance contained in real-world data into different components, which arise from different sources of spectral changes

spectrometer in transmission or in attenuated total reflection (ATR) mode; ATR has become one of the most popular methods in the spectroscopy of biological materials [123]. The easily perturbed infrared spectra contain huge amounts of information, which can be converted into a picture of the sequential response of the protein to the applied external perturbation. However, this picture can be blurred by various imperfections, distortions, and noise, which are present in the routine measurements. The picture can even be artificially enhanced by excess preprocessing. The quality of information retrieved from 2DCoS spectra largely depends on both the quality of the raw spectra and the modification of the spectra at different correction and processing steps. This dependence is a general problem that also concerns other multivariate methods of analysis; therefore, knowledge of good practices in infrared data pretreatment is necessary. Some useful guidelines can be found in [124].

As the scheme in Fig. 8.2 shows, the total variance contained in the raw data must be subjected to procedures that allow efficient elimination of the error caused by common effects, such as noise or baseline fluctuations. After the successful removal of these unwanted components, some additional procedures could be necessary depending on the conditions of the measurements. Whenever aqueous protein solutions are analyzed, it is critical to remove the strong contribution from the water bands. In ATR experiments, effects related to the refractive index change both in the

absorption band and in the series of spectra, and effects related to the wavelength dependence of the penetration depth [125] should be reduced.

Usually, in the final stage of the pretreatment process, the spectra must be normalized. This procedure is necessary if there is a strong contribution from the nonspecific variance and a weak contribution from the specific variance. Skipping this step could lead to a loss of information from the synchronous spectra and the false interpretation of the asynchronous spectra. Normalization allows one to isolate frequency-specific spectral variations arising from secondary structure changes from those nonspecific spectral variations shaped by ordinary effects such as density or concentration changes. Finally, when the well-separated specific changes are processed by 2DCoS, the changes can be converted into the different structural events that shape the folding funnel. One of the most serious problems in 2D correlation spectroscopy is the appearance of artifacts in the 2D maps; these artifacts could result from pretreatment steps that are incorrectly performed or omitted and from improperly selected normalization [84].

8.3.2 External Perturbations

There are numerous physical and chemical perturbations that have been used to induce a specific spectral response, which is the unique answer of the different structural elements of the biomolecules to the applied stimulation. The most popular perturbations in protein studies are temperature, H/D exchange, pressure, protein concentration, pH, and chemical denaturants.

8.3.2.1 Temperature

In the field of protein research through 2DCoS, temperature is the most commonly used static perturbation. Temperature changes are commonly used because ordinary changes in density, which give rise to nonspecific spectral changes, usually are accompanied by structural changes in the protein backbone upon heating or cooling. For a range of amide bands, these changes are manifested by specific spectral variations that are well suited to the 2DCoS analysis. One of the first such studies concerned the investigation of the order of intensity changes and the prediction of the sequence of structural evolution during the temperature-induced unfolding processes of the ribonuclease A (RNase A) and λ Cro-V55C repressor proteins [38, 50]. The aim of the research was to determine whether the folding/unfolding process is fully cooperative or whether the process proceeds through sequential elements. To address this question, it was necessary to use tools that could resolve the amide I envelope into components that could be assigned to different events. Being aware of the facts presented in [4], Mantsch's group has turned to the 2DCoS approach, which provides a perfect method to improve the resolution of the individual band components contributing to the amide I envelope and which facilitates the detection

of a potential sequence of events in response to changes induced by a temperature increase from 20 °C to 95 °C. Using the procedure proposed by Noda [1, 6] for the interpretation of the synchronous/asynchronous changes, three spectroscopic events have been suggested for the RNase A system on this temperature scale; these events involve structures that contribute to the following peaks: 1,637/1,692 cm^{-1} (step 1), 1,656 cm^{-1} (step 2), and 1,644 cm^{-1} (step 3). Based on the well-defined relationship between frequency and secondary structure [126], the spectral changes has been "translated" into the following sequence: in the first step of the process, the partial unfolding of the more strongly bonded β-sheet fragment occurred. In the next steps, the α-helices and the weaker parts of the β-sheet unfold [38].

The thermally induced unfolding of RNase A has also been investigated under reducing conditions maintained by the presence of 2-mercaptoethanol [83]. In these studies fluorescence spectroscopy has also been employed, and 2DCoS was combined with principal component analysis (PCA) to interpret the infrared data. The paper was published in a special volume of Vibrational Spectroscopy dedicated to Professor Heinz Siesler, a pioneer in NIR spectroscopy of polymers and applications of 2DCoS for analysis.

Arrondo and coworkers have also frequently applied temperature as a perturbation mode. From temperature-perturbed infrared spectra of rat liver methionine adenosyltransferase (MAT), the secondary structure and the events occurring during the thermal unfolding of MAT were characterized [62]. First, based on the thermal activity profiles and one-dimensional analysis of the infrared and fluorescence spectra, the unfolding and aggregation process of MAT was divided into three temperature stages. Next, analysis of the 2D correlation maps independently obtained for each region indicated that the α-helix and β-turns change first. Aggregation, which occurred at the expense of the α-helix, β-sheet, and the residual β-turns, a major event in the second stage. Correlation maps corresponding to the highest temperature range indicated no further major changes in protein conformation and proved that denaturation is complete in the last stage.

In [65], urea and thermal treatments were combined to study the effect of urea on the thermal denaturation of an integral membrane protein (sarcoplasmic reticulum Ca^{2+}-ATPase) that pumps calcium out of the cytoplasm during the relaxation of striated muscle. The action of the chaotropic agent urea on the thermal stability of this protein was examined by temperature-perturbed 2DCoS for the control sample (non-urea treated) and 3 M and 4 M solutions of urea. To check if the urea-induced changes were reversible or non-reversible in character, the two urea systems were washed twice with urea-free buffer and were heated; the obtained 2DCoS maps were compared with those prior to washing and with the control sample. 2D-IR-CoS spectroscopy revealed that in the 3 M system, the urea perturbed the α and β structures and stopped the thermal transition in a reversible manner. However, in the 4 M urea system, the protein structure was changed in a non-reversible way.

Temperature perturbation has also been used in studies of the influence of calcium on the stability of chicken annexin A5 (cA5), which can reversibly bind to fatty acid phospholipid-rich membranes in the presence of calcium [59]. In temperature-dependent synchronous 2D correlation maps of cA5 and two mutants of

cA5 measured in the absence or presence of Ca^{2+}ion, two populations of α-helical structures have been identified, i.e., the canonical α-helices (1,650 cm^{-1}) and solvated α-helices with those α-helices in helix-helix interactions (1,634 cm^{-1}). The results strongly support results obtained from circular dichroism and differential scanning calorimetry; these results reveal more compact tertiary structures, which have higher thermal stability, after calcium binding.

The formation of amyloid fibrils from the native state of the activation domain of human procarboxypeptidase A2 (ADA2h) and three mutants with different aggregation propensities was studied by Cerdà-Costa et al. with temperature as the perturbing agent [81]. A sample at pD 3 was heated from 20 to 90 °C in 10 min to trigger the aggregation process. Next, IR spectra were measured every min for 30 min while the sample temperature was maintained at 90 °C. Unlike the studies described above, in this case, temperature was not the direct perturbation. The temperature change initiated the time-evolving process of fibril formation, which was the direct perturbation. Inspection of the two distinct autopeaks at 1,616 cm^{-1} and centered around 1,647 cm^{-1}, which are associated with the amyloid fibril and partially unfolded α-helices, respectively, allowed for a detailed analysis of the conversion of the secondary structure of ADA2h from the initial monomer to the amyloid fibril.

The sequential unfolding events of horse, cow, and tuna ferricytochrome c (fcyt c) were determined, and whether the thermal unfolding processes for the structurally homologous proteins proceeded by the same steps was examined [53]. 2DCoS has helped to find subtle but significant differences in the order of the thermal unfolding events for the three proteins, despite their high sequence identity (>97 %). In cow fcyt c, the unfolding of all of the helical components occurred cooperatively, but in horse and tuna fcyt c, the helical components behaved as subdomains that unfolded separately.

A few papers were published by Yan et al. in which the temperature-perturbed infrared spectra of myoglobin [61], hemoglobin [64, 69], and bovine pancreatic ribonuclease A [74] were analyzed by 2DCoS to describe the heat-induced unfolding and aggregation processes for the three systems. The phenomenon of protein aggregation attracts many researchers because this process is strictly correlated with protein fibrillation [127]. Although a two-state model for thermal denaturation of myoglobin arose from traditional 1D analysis, subtle noncooperative conformational changes identified at low temperatures by 2DCoS forced the postulate that myoglobin aggregates under conditions when native structures dominate [61]. The appearance of an aggregated β-structure before the unfolding of α-helices was surprising at that time. In contrast, the sequence of events from the 2D analysis of the temperature-dependent spectra of hemoglobin clearly indicated a two-stage thermal transition [64]. The aggregation process is preceded by the exposure of the helical structures to the solvent; next, from 44–54 °C aggregated forms evolved from the exposed elements. The process is completed at 70 °C by the unfolding of structures deeply buried in the hydrophobic interior. 2DCoS analysis of the thermal denaturation of RNase A in acidic (pD 6.0) and basic (pD 8.0) conditions [74] has revealed that the aggregation was closely associated with the oligomerization of RNase A and that both processes proceeded at high temperatures. The sequence of events for

RNase A has also indicated that structural rearrangement most likely initiated the aggregation process, as was observed for hemoglobin [63]. 2DCoS studies of the aggregation of hemoglobin and RNase A did not confirm the then surprising results obtained for myoglobin [61].

The aggregation process induced by temperature and by pH has been studied in whole porcine plasma (serum albumin + globulins + fibrinogen), serum (serum albumin + globulins), and its fractions (serum, albumin and globulins) [82, 88]. Infrared results aided by 2DCoS have shown that the behavior of serum albumin and globulins during the thermal treatment has been strongly changed in the presence of fibrinogen. Despite the fibrinogen-correlated differences, in all the systems, the α-helical structure has been the most temperature-sensitive conformation. The disappearance of the α-helical structure and formation of the heat-induced intermolecular β-sheet structure proceeded in a closely related way.

Temperature could be a very advantageous perturbation in studies of the interactions between intrinsic and extrinsic proteins and lipid bilayers because during the heating process, lipid bilayers undergo a characteristic phase transition from a gel to a fluid state; this transition could act as an additional internal perturbation of the protein structure. Such 2DCoS-aided investigations were performed for the first time by Lefèvre and Subirade [54] who studied the interactions between β-lactoglobulin (β-lg) and dimyristoylphosphatidylglycerol (DMPG) bilayers at pD above (pD 7.4) and below (pD 4.4) the isoelectric point of the protein (pI = 5.2). The spectra were measured for β-lg alone and in the presence of DMPG as a function of temperature over the range of the main transition of DMPG to determine whether the conformational modifications of the protein in the presence of lipids were caused by thermal effects or whether the conformations were modulated by the chain-melting phase transition. To answer this question, synchronous and asynchronous 2D correlation maps were constructed from the spectral variations in the two frequency ranges, i.e., 3,000–2,800 cm^{-1} and 1,685–1,600 cm^{-1}. The heat-induced structural changes of β-lg did not correlate with the conformational modifications of the phospholipids along the phase transition. The thermal behavior of the protein was not modulated by the lipid chain melting; instead, the behavior represented the heat-induced protein rearrangement in the presence of DMPG. Moreover, analysis of the sequence of events during the thermal unfolding of β-lg at pD 4.4 and 7.4 has shown similar initial structures and similar subsequent thermal effects under the conditions of the two pD values.

The influence of temperature and time on the denaturation and aggregation phenomena of cytochrome c (cyt c) in the presence of DMPG was investigated in [56]. Aggregation at 65 °C has been confirmed by a band at 1,616 cm^{-1}, which is an IR signature of this process. The studies assisted by 2DCoS have revealed that aggregation starts between nearly native proteins whose unfolding leads to further aggregation of cyt c. In a later stage of this process, the formation of intermolecular bonds and the unfolding of the α-helices appear to be simultaneous. Based on this finding, a two-step aggregation process was proposed. 2DCoS results obtained from data collected during the heating period have shown a loosening of the tertiary

structure, which preceded the aggregation process. Additionally, an increased number of β-turns at the expense of β-sheet content was reported. That study confirmed the potential of 2D-IR-CoS spectroscopy to investigate the intrinsic temperature-mediated or direct lipid-mediated change in protein conformation in a lipid bilayer environment.

It is outside of the scope of this article to mention all of the results in which temperature perturbation has been applied in 2D infrared correlation spectroscopy. The selected examples are intended to illustrate that there is only one step between the very common near-IR/mid-IR temperature-dependent experiment and 2D correlation spectroscopy. Why should this step be taken? A clear answer arises from the above examples: In contrast to ordinary one-dimensional analysis, two-dimensional infrared correlation spectroscopy can provide insight at the molecular level into the underlying mechanisms responsible for protein unfolding, including protein aggregation and fibrillation. Can we expect anything more from a routine temperature-dependent IR experiment?

8.3.3 Concentration of protein

Concentration has been a commonly used perturbation in investigations by Ozaki and co-workers. Variations in the concentration of protein at a given temperature have been the main source of spectral changes subjected to 2DCoS analysis. If all of the changes exclusively follow the Lambert–Beer law, the spectra should not exhibit any asynchronous properties. However, well-observed asynchronous peaks have shown specific spectral changes caused by an increase in ovalbumin concentration from 2 to 8% by weight (wt) [90]. These changes were a good spectral signature of the structural changes caused by the water-mediated protein–protein intermolecular interaction. Analysis of the synchronous and asynchronous maps generated from the concentration-dependent spectra measured in a broad temperature range has allowed the concentration-controlled interaction to be monitored as the protein structure evolved from the native to the denatured unfolded state. The use of near-IR spectra in a range of overtone and combination bands attributed to water and protein, respectively, has made it possible to analyze the role of hydration water in the protein–protein intermolecular interactions at different stages of thermal denaturation. It has been shown that the hydration of ovalbumin is almost unchanged from 45 to 67 °C; in this range ovalbumin molecules were in a natively folded state. A sudden change in the hydration was observed in the narrow temperature range of 67–69 °C, in which the unfolding of the ordered secondary structures started. The hydration again remained nearly unchanged upon further heating to 80 °C.

The nonspecific spectral changes induced by concentration are rather weak compared with those generated by other types of perturbations. Thus, especially in the case of IR-ATR measurements, it should be checked the influence of the interaction of the protein with the surface of the internal reflection element (IRE); this interac-

tion could lead to an adsorption process. According to [127], this process may be responsible for substantial structural distortions in the secondary structure and for additional spectral variations that cannot be eliminated through any mathematical pretreatment procedure. Therefore, it was interesting to compare the adsorption-induced changes with the changes caused by concentration [47]. To answer this question, power spectra that present the overall extent of the intensity changes caused by the given perturbation process have been used. A comparative analysis of these power spectra has shown that the concentration changes were 10 times more intense from those arising from adsorption for the example of β-lactoglobulin measured as a function of concentration (from 1 to 5% wt) and time for a 5% wt system. This result has convinced us that concentration was the dominant perturbation and that 2D results can be discussed while the minor contributions induced by adsorption are neglected.

Similar experimental protocols have been used to study the thermal denaturation of human serum albumin [92] and the molten globule-like state of ovalbumin at acidic pH values [95]. In these analyses, the advantages of power and slice spectra in monitoring the process of protein unfolding have been demonstrated. In [92], a plot of the slice spectra at different temperatures extracted from concentration-perturbed asynchronous contour maps at 4,600 cm^{-1}, the amide B–amide II combination transition, has clearly indicated a sudden change in the secondary structure near 60 °C. A similar plot for the band near 7,000 cm^{-1}, assigned to a combination mode of water, has confirmed the active role of water in the process. The change from near-IR to mid-IR radiation has allowed the study of the roles of different secondary structures and side chains in the transition [60]. The asynchronous spectrum generated from the concentration-dependent spectral variations at a temperature before the transition has shown that the exposed part of protein composed from the side chains, random coil, and extended chains has been more sensitive to the concentration changes than the α-helices and α-turns hidden in the protein core. However, the identical analysis of the concentration-dependent spectra collected at a temperature above the transition has revealed the opening of the compact core because the changes assigned to β-turns were identified before those in the extended chain, random coil structures, and α-helices.

Concentration has also been used as an external perturbation by Kamerzell et al. [98], who have investigate the interactions that modulate the nonideal behavior of two immunoglobulins (IgG) in highly concentrated solutions. With the same data pretreatment as in [47], the nonspecific spectral changes were minimized, and the remaining changes that deviated from linearity with concentration increase from 20 to 120 mg/mL were subjected to 2DCoS. The sequence of spectral events as a function of increased concentration of IgG has revealed that the variations in the β-sheet and turn regions have occurred before the intensity variations in the disordered and α-helical regions. Moreover, the 2DCoS results supported by PCA have shown that Glu and Asp residues triggered the processes that resulted in association of IgG. In summary, increased intermolecular hydrogen bonding interactions and electrostatic interaction modulate the intermolecular association and the nonideal behavior.

8.3.4 H/D exchange

Hydrogen–deuterium (H/D) exchange of the amide protons has found important applications in many experimental studies of protein structure. The contributions of different secondary structures to the amide bands can be efficiently separated because of the different exchange rates, which are controlled by the solvent accessibility and hydrogen bond stability of the amide protons [129]. Infrared measurements as a function of time collected a large data set dominated by specific spectral changes in the range of the amide I, amide I', and amide II' bands; these changes make H/D exchange an excellent external perturbant.

The advantage of this perturbation in the 2DCoS analysis was first shown in Pézolet's group for an aqueous solution of myoglobin measured by IR-ATR spectroscopy [36]. Because of the different rates of the H/D exchange for different elements of the secondary structure of myoglobin, this method has allowed the identification of at least five components of the amide I band; these components were assigned to the α-helix, intermolecular β-sheet, β-turn, and random coil conformations. These results have initiated many investigations in which H/D exchange has been chosen as a perturbation. This methodology has been applied to monitor the changes in secondary structure of a small basic protein found in the endosperm of wheat seeds (β-purothionin) because of the incorporation of this protein into lipid bilayers composed of dimyristoylphosphatidylglycerol (DMPG). Comparative analysis of the 2DCoS synchronous and asynchronous maps allowed the observation of an increase in the α-helix content of β-purothionin; this increase has confirmed the formation of a functional protein channel in the membrane upon the binding of the protein to the lipid membrane [67].

This effective method of unraveling the different components in the poorly resolved amide I, II, and II′ bands of proteins was also used by Meskers et al. [4443] to determine whether the H/D exchange process in streptavidin was modified in the presence of biotin. Generally, the binding of a ligand can be accompanied by conformational changes at the protein site that can propagate throughout the entire protein. Therefore, the analysis was focused on distinguishing specific absorbance variations in the 1,400–1,700 cm^{-1} range and on the assignment of these variations to subtle differences within the secondary structure. 2DCoS results have revealed that the strong non-covalent interactions between streptavidin and the biotin ligand inhibited the exchange for some of the amide protons, which were part of the β-sheet structure of this protein. This inhibition was identified from the slowly proceeding changes at 1,530 and 1,445 cm^{-1}, which were assigned to the amide II and II′ frequencies of the β-sheet.

Wu et al. [57] used 2D-IR-CoS supported by PCA to investigate the kinetics of H/D exchange in human serum albumin (HSA). From the PCA results, the exchange process has been separated into three stages that have different structural change profiles, as shown by the loadings. The slice and power spectra extracted from the synchronous and asynchronous maps were independently calculated for the data grouped into the three time domains. The results have shown that during

first stage, i.e., within the initial 6 min, amide protons from the structures most exposed to solvent, the outer surface of HAS, are exchanged for deuterons. This process was identified from the changes at approximately 1,683 cm^{-1}, which were assigned to the β-turns. During the next stage, between 6 and 14 min, because of the penetration of D_2O deeper into the more compact interior of the HSA, protons from the β-turns that were initially less solvent accessible are exchanged and give rise to the peak at ~1,666 cm^{-1}. In the last part of the H/D exchange, the peak observed at 1,653 cm^{-1} shows marked proton replacement in the helical part that forms 67% of the total structure of HSA. It was estimated that approximately 25% of the α-helix structures in HSA are unaffected by the H/D exchange even 4 hours after the beginning of the experiment.

The time evolution of the spectral variations induced in the amide I and amide I' range by H-D exchange at different pH values has been applied to examine the flexibility of immunoglobulin; this flexibility is essential for the function of IgG [76]. In the analysis, the heterocorrelation variant was also employed to confirm the correlations between two separate data sets measured at different pH values. The H/D perturbation-based 2DCoS results have delivered evidence that IgG is composed of multiple regions with differential flexibility and conformational heterogeneity; these regions were differentially perturbed by solution conditions such as the pH. The H/D-perturbed 2DCoS has allowed for an improved understanding of the relationships between IgG dynamics and stability.

The secondary and tertiary structures of the truncated protein covering the first 120 aa of the hepatitis C virus core protein (HCVc-120), which is assembled in vitro into nucleocapsid-like particles, have also been studied through the dynamics of H/D exchange [77]. The infrared spectra measured as a function of the H/D exchange time, over a period of two hours, were examined by both principal component analysis and two-dimensional correlation spectroscopy. The PCA results were a good basis for the selection of two time domains for further two-dimensional correlation analysis. For PC1, the spectra collected over the initial 10 min were clearly distinct from those measured between 13 and 120 min after the initiation of the H/D isotopic exchange. The temporal characteristics of the H/D exchange obtained from the 2D analysis performed separately for the two sets of data has shown that in most arginine residues, protons from side-chain fragment were exchanged within the first 10 min. This finding has allowed the postulate that the positively charged Arg residues, which are easily accessible to aqueous environment, were prone to accept negatively charged phosphate groups from the viral RNA during the assembly of the HCV; this process should have led to an increase in the thermodynamic stability of the nucleocapsid. The fast H/D exchange process for the arginine residues were followed by changes in the β-turns and unordered structures and were finally followed by rearrangements in the β-sheets, which might have linked the protein monomers.

H/D exchange supported by simulation of this process has been a common procedure used by Pastrana-Rios's group. For example, this approach has allowed the resolution of the dynamics of the molecular events that occur upon H/D exchange in two of eight different epidermal growth factor receptor (EGFR) ligands (TGF-α and EGF) [71]. The studies were motivated by the unfortunate large overlap be-

tween bands arising from vibrations of exchangeable protons from N–H amide bonds and protons belonging to the Arg and Tyr side chains. In the experiment, subsequent spectra were acquired every 10 min for approximately 14 h, and the simulated spectra that mimicked the changes of the experimental spectra as much as possible were subjected to 2DCoS. The comparison of only the synchronous spectra for the experimental and simulated data from 1,800–1,400 cm^{-1} resulted in the separation of the highly overlapped absorbances from the backbone and from the Arg and Tyr side chains. In [78], the studies were extended to next three EGFR ligands (ER, HB-EGF, AR), and in the analysis, the asynchronous spectra were also employed; these spectra allowed the detection of the differences in the exchange rates for these five ligands, within the side chains and the main back-bone; these differences may be critical for receptor regulation. Other types of systems for which very detailed assignments of different structural elements to sequential changes in the H/D exchange have been presented are centrin [66] and the complex of centrin with melittin [85].

8.3.5 Pressure

Smeller and Heremans [45] first applied 2DCoS to analyze the pressure-induced changes in the FTIR spectra of proteins. Initially, a small protein such as bovine pancreatic trypsin inhibitor (BPTI, 8 kDa) was subjected to H/D exchange overnight at room temperature. Next, a high pressure (0–1 GPa) was applied as a perturbation, and the FTIR spectra were measured. The application of 2DCoS has been very efficient in the separation of the spectral features arising from H/D exchange from those spectral features that are pressure-induced. In the low pressure region (0–0.3 GPa), both conformational and exchange effects took place whereas above 0.5 GPa (up to 0.9 GPa), the exchange was complete and only conformational changes were observed [46]. The next problem of interest was the comparison of this small protein with lipoxygenase (LOX), which is a large protein (102 kDa). The application of 2DCoS helped unravel the relation between the conformational changes and the kinetics of the H/D exchange processes for the two types of proteins [51]. The pressure range was divided into three parts for the 2D correlation analysis: the low pressure range (2–4 kbar), the transition region (4.7–9.0 kbar), and the upper pressure region (8.5–10.6 kbar). It has been observed that for BPTI, the low pressure regime acts preferentially on the conformational dynamics and promotes H/D exchange. For LOX, the pressure induced irreversible conformational changes, which were responsible for increased exposure of the interior parts of the protein to the solvent.

Poly(L-lysine) (PLL) is a very attractive polypeptide because under different conditions, it can form one of three stable conformations: an α-helix, an anti-parallel β-sheet, or a random coil. The pressure resistance of 70 kDa PLL has been checked over a pressure increase from 0.1 to 1,150 MPa [87]. Guided by PCA results, the pressure dependent spectra have been separated into low (0.1–400 MPa)

and high (700–1,150 MPa) pressure ranges and have been independently analyzed by 2DCoS. The obtained results show increased hydration of the α-helix conformer (low range) and the gradual transition to α′-helix (high range). In contrast, the initial β-sheets and random coils are first transformed to an α-helix at pressures < 400 MPa and are next transformed to an α′-helix in the high pressure range.

More recently, Wu's group has investigated the pressure-induced structural changes in isolated bovine serum albumin and bovine serum albumin used as a protecting ligand for gold nanoclusters, AuNCs@BSA. In these studies, both fluorescence and IR spectroscopy assisted by 2DCoS were involved [89]. The 2DCoS results have supported the results that under compression, more buried α-helical structures of BSA are lost whereas in the AuNCs@BSA, a more flexible conformational state is formed at the boundary of the gold nanoclusters.

Dzwolak et al. wrote a review of the application of high pressure in studies of protein by FTIR spectroscopy [130]. The authors have also used the 2DCoS analysis to confirm the acceleration of the kinetics of H/D-exchange with temperature upon the heating of α-lactalbumin (α-la) from 20 °C to 42.5 °C and upon a pressure increase from 0.1 MPa to 180 MPa [48]. Comparison of the synchronous and asynchronous maps obtained for the two accelerated systems with the maps of the control system (the time-evolving H/D exchange) has shown that the pressure and temperature treatments differently changed the kinetics of the H/D exchange. The comparison has revealed that the compactness inside the different secondary structures of α-la is modified differently by the applied pressure and temperature.

8.3.6 pH

Modulation of pH induces specific conformational changes that are relevant for protein function; moreover, the response of the protein structure to pH leads to many well-defined spectral events [131]. Therefore, experiments have frequently been conducted for samples at different pH values [54, 74, 76, 77, 81]; however, during the perturbation period, the pH of the system was maintained at a constant level. Processes accompanying the protonation/deprotonation of the basic and acidic side chain groups result in reorganization of the tertiary and secondary protein packing and lead to changes in the absorbance; because of their high specificity, these changes are especially well suited for 2DCoS analysis. Using 2DCoS to pH-perturbed mid-IR spectra of human serum albumin (HSA), Murayama et al. [55] have tried to find answers to following questions: in which direction is the secondary structure modified? and what is the role of microenvironmental changes at the side chain sites when HSA is transformed between two isomeric forms as the pH decreases from pH 5.0 to 3.0? The synchronous and asynchronous spectra have been calculated from the spectral changes and have been grouped into three pH ranges corresponding to the N isomeric form (pH 5.0–4.4), the N–F transition (pH 4.6–3.8), and the F isomeric form (pH 3.8–3.0). Because of the powerful deconvolution ability of 2DCoS, four sub-bands were detected at 1,740, 1,715, 1,705, and

1,696 cm^{-1}, which arise from the C=O stretching modes of the free COOH groups and the COOH groups involved in hydrogen bonds of increasing strength. This analysis strongly supported the hypothesis that protonation of the COO^- groups of the glutamic and aspartic acid residues has occurred prior to the structural changes that develop in the α-helices and that dominate in HSA folding. The N–F transition has been correlated with the presence of free COOH groups and the formation of medium-strength hydrogen bonds between the COOH groups and groups from the main chain, side chains or water molecules. The asynchronous spectrum confirmed that the hydrogen bonds play a key role in the initiation of the N–F transition and trigger the α-helical changes specific to the transition. The synchronous and asynchronous spectra of the F form have shown that the β-strands and β-turns of HSA change significantly in the pH region between 3.8 and 3.0.

2DCoS has been applied to pH-perturbed mid-IR spectra of porcine plasma proteins [82]. The system has a variety of proteins with diverse secondary structure, and when subjected to a pH decrease from 7.5 to 4.5, the proteins undergo both unfolding and aggregation processes at 30 °C. A detailed analysis of the sequence of spectral intensity changes that take place during the pH decrease has been performed for six peaks assigned to the most common secondary structure elements of the proteins present in the plasma. The studies have shown that in the denaturation process that begins at a higher pH, the globulin fraction is primarily involved whereas the serum albumin fraction is denatured under more acidic conditions. pH perturbation was also used in the first NMR-Raman heterospectral correlation studies of the structural evolution of silk fibroin [104].

8.3.7 Other Types of Perturbations

The properties of silk fiber from *Bombyx mori* have been analyzed by dynamic step-scan FTIR measurements as a function of a sinusoidal mechanical strain [52]. The dynamic spectra revealed that the stress-induced dynamic reorientation in the fibroin film was mainly localized in the segment with a β-sheet conformation and was almost synchronous with the applied mechanical strain. Increased surface pressure has been a common perturbation in studies conducted by Dluhy's group, and infrared reflection–absorbance spectroscopy was used for the measurements [35, 58, 68].

2DCoS was suitable for the real-time detection of the molecular processes occurring in an intact retina as a response to light [79]. 2DCoS has facilitated the identification of bands arising from the same molecular chromophores even in the presence of very complex spectral changes. A process in which a β-sheet-rich protein is released from the disk membrane and is redistributed from the outer segment towards the inner segment of rod cells has been postulated.

A varying concentration of denaturing agent has been used as a perturbation in the analysis of the influence of bromoethanol on the structure of β-lactoglobulin studied by IR in both ranges [37]. Two different pictures of sequential changes

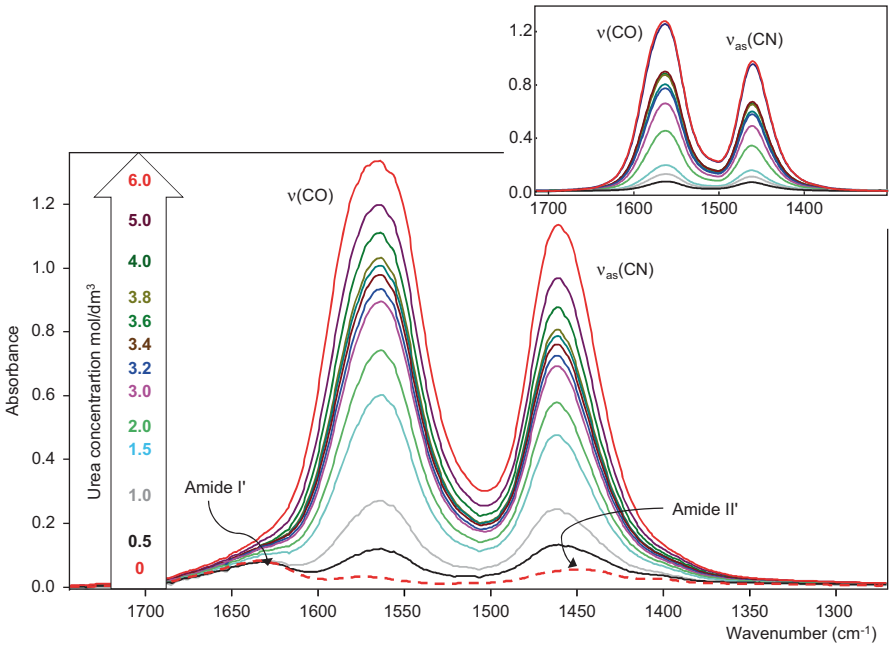


Fig. 8.3 Evolution of the absorbance from 1,700 to 1,300 cm^{-1} as a function of the urea concentration for the ternary system after pretreatment procedures. The spectrum marked with a dashed line is a solution of β-lg (2 wt %) without urea. The inset shows the same types of results for the binary system. (Reproduced with permission from Ref. No. 30. Copyright (2009) American Chemical Society)

obtained from the two measurements have confirmed the high sensitivity of the amide I band to the changes in the secondary structure and the low sensitivity of the amide B/amide II combination band located in the near-IR range to the changes. More information has been found in the C–H stretching/bending combination bands; the properties of these bands correlated with the interaction between bromoethanol and β-lg during the transition of β-lg from β-sheet to α-helix conformation by two parallel pathways.

The next example comes from our studies focused on determining the role of urea in the unfolding of β-lg induced by an increased concentration of urea [30]. This example did not use the typical procedure based on the analysis of the amide bands because the bands are strongly overlapped by intense bands from urea and subtraction of the urea contribution from the total absorbance changes was not allowed. Therefore, the analysis was concentrated on the two bands of urea, i.e., the C=O stretch and the asymmetric CN stretch; the variations with concentration changes are shown in Fig. 8.3. To obtain a detailed view of the many-layered problem of the urea–protein interaction, the changes in the ternary system (β-lg–urea–water) have been compared with those observed in a binary urea–water system [132]. A simple analysis based only on the comparison of the relative absorbance

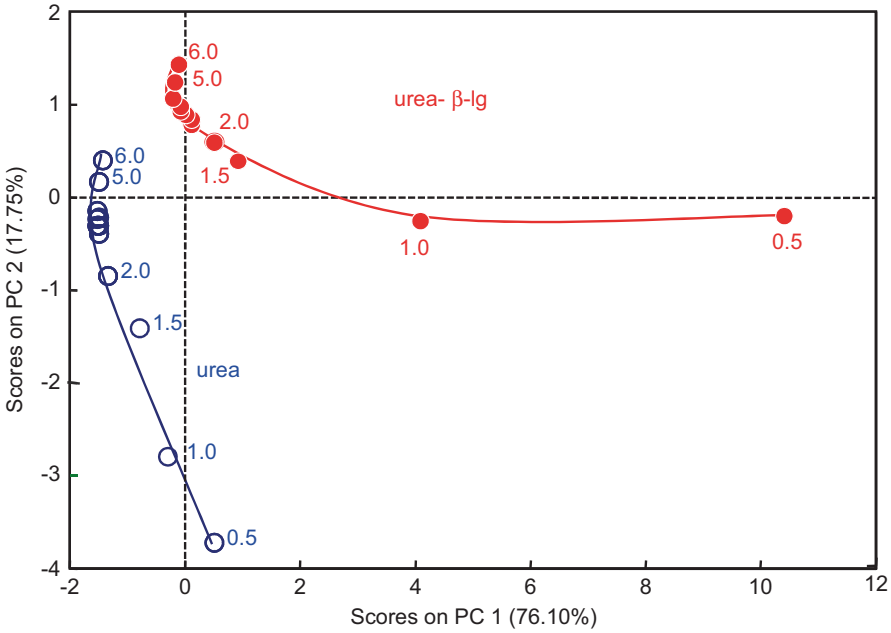


Fig. 8.4 Plot of the scores for the first two principal components for the spectra of the system with (●) and without (○) β-lg combined into one matrix. (Reproduced with permission from Ref. No. 30. Copyright (2009) American Chemical Society)

changes for the two bands has indicated the concentration-controlled involvement of the C=O groups in the interactions with the β-lg molecules. When the binary and ternary sets of spectra were combined into one matrix and subjected to PCA, more details about the differences in the urea–water interaction caused by β-lg were obtained. The resulting plot of the scores (PC1 vs. PC2) from this PCA analysis are presented in Fig. 8.4 and have confirmed that the concentration-dependent association phenomenon of urea in aqueous solution has been strongly influenced by the interaction with β-lg, mainly at concentrations lower than 2 M for which urea was involved in ribbon and chain dimers [132]. The changes were captured by PC1, which explained more than 76% of the total absorbance variations. PC2 described almost 18% of the variations directly related to the absorbance changes that arose from urea-induced unfolding of β-lg; this unfolding mainly occurs at concentrations above 2 M. This classification of the changes has been confirmed by the plot of the loadings for the two components.

Figure 8.5 shows the asynchronous spectrum calculated for the ternary system only. To more easily detect the 2D peaks, we used a slice spectrum at 1,563 cm^{-1}, which is shown on top. Based on the order of absorbance changes determined from the asynchronous spectrum with increased concentration of urea (1,563 cm^{-1} → 1,545 cm^{-1} ≈ 1,590 cm^{-1}), the following scenario for the interactions in the tertiary system was postulated. The process begins with a direct water–β-lg interaction,

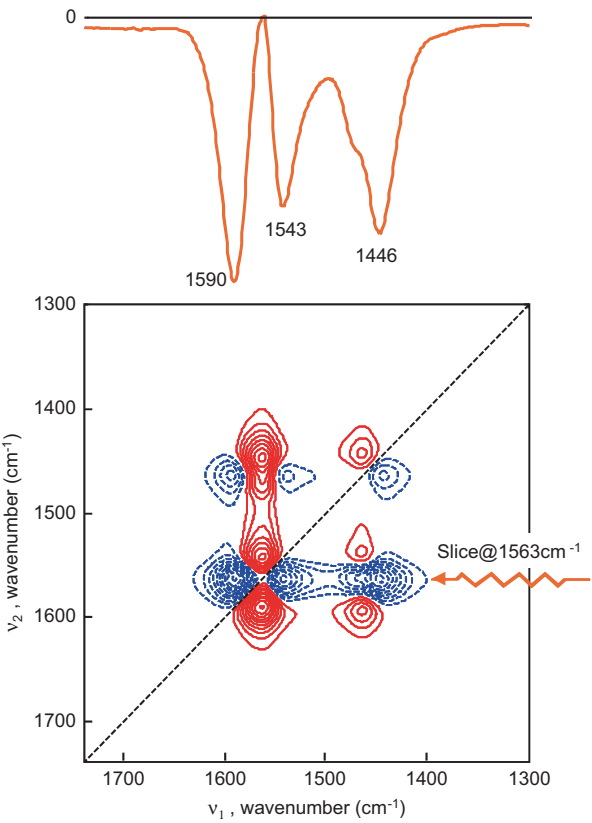


Fig. 8.5 Homosample asynchronous 2D correlation spectrum for the ternary system and the slice spectrum extracted at the position of maximal asynchronicity presented above. The solid and dashed lines represent the positive and negative cross peaks, respectively. (Reproduced with permission from Ref. No. 30. Copyright (2009) American Chemical Society)

which affects the interaction between urea dimers and water molecules and leads to the variations at 1,563 cm^{-1}. As the urea concentration increases to 2 M, the structure of the water changes from the interaction with the urea dimers. Consequently, the water–β-lg interaction weakens and the hydrophobic forces that hold the β-lg molecules in the native state are reduced. During this water-mediated stage, the secondary structure of β-lg is not modified, as shown by the PCA results. At the next stage, above 3 M, the denaturation mechanism has changed into the direct, urea-mediated process. The expanded structure of β-lg opens through direct interaction with urea associates, and most of the C=O groups are hydrogen bonded, but some of them are also free. Because they experience different influences from the polar and nonpolar fragments of the β-lg chain, the two peaks observed at 1,543 cm^{-1} and 1,590 cm^{-1} are assigned to bonded and nonbonded C=O groups [132], respectively.

The heterosample asynchronous spectrum shown in Fig. 8.6 has allowed us to confirm that the properties of the urea associates are affected by the presence of β-lg. The strong asynchronous peaks at 1,563 and 1,562 cm^{-1} assigned to the hydrogen-bonded C=O groups in the ternary and binary systems, respectively, confirms that the changes proceed with different kinetics. Based on sign of these peaks, it has

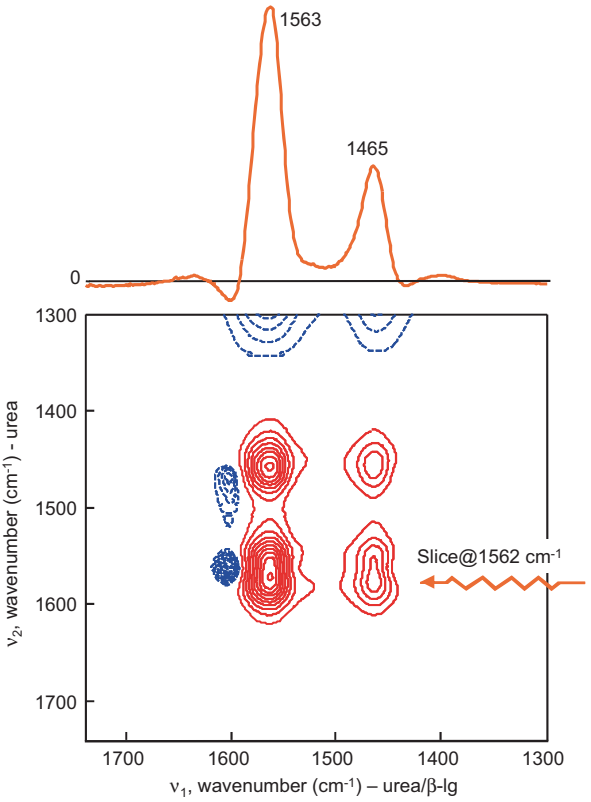


Fig. 8.6 Heterosample asynchronous 2D correlation spectrum for the two systems and the slice spectrum extracted at the position of their maximal asynchronicity shown above. The solid and dashed lines represent positive and negative cross peaks, respectively. (Reproduced with permission from Ref. No. 30. Copyright (2009) American Chemical Society)

been postulated that the presence of β-lg makes the urea dimers less stable and more sensitive to changes in the urea concentration.

8.4 Developments in Generalized 2DCoS Useful in Protein Research

8.4.1 Heterospectral Correlation

The summary plot in Fig. 8.7 shows the various optical spectroscopies that have been combined with 2DCoS to provide complimentary information about protein structure at different levels of organization and interaction with environment. Such a large range of methods involved in protein studies has opened a new possibility for 2DCoS called heterospectral correlation analysis. Heterospectral correlation analysis involves comparison of the spectral changes induced by identical perturbations and measured by closely related spectroscopic methods. This approach can also be involved in combined analyses of the results obtained from completely dif-

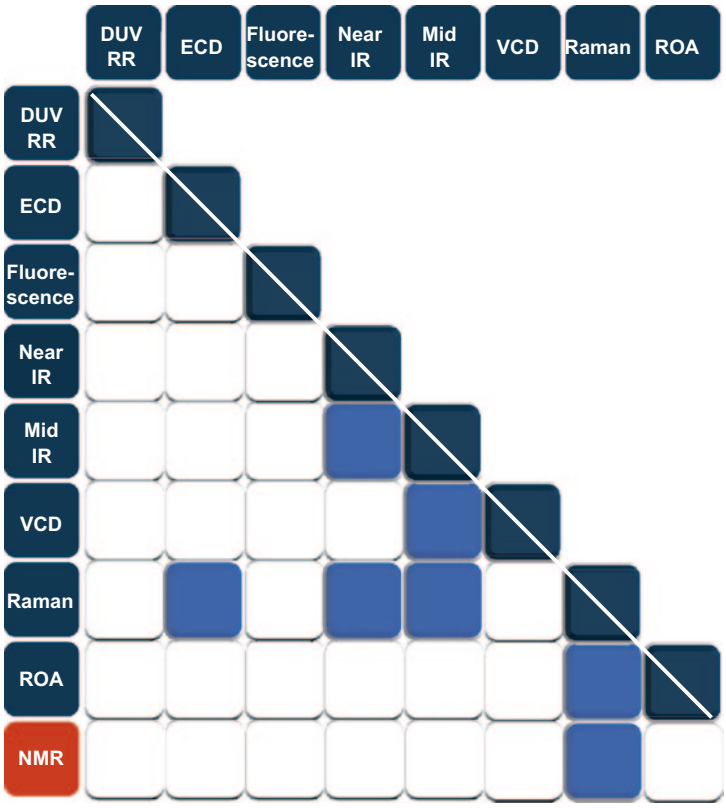


Fig. 8.7 Summary of the optical spectroscopy methods that have been applied in the homo- ■ and heterocorrelation ■ modes of 2DCoS. NMR as a non-optical technique is marked in red

ferent types of spectroscopies. The blue off-diagonal boxes in Fig. 8.7 correspond to methods that have already been applied in heterospectral correlation analysis, whereas the dark blue diagonal boxes denote homospectral correlation analysis. The light grey boxes show new possibilities for heterocorrelation analysis that have not yet been realized.

The earliest attempt at this type of analysis was performed by Barton et al. [133]. The researchers used the near-IR ↔ mid-IR statistical 2D correlation to analyze forage. The analysis has allowed the correlation among bands in two different types of spectroscopies to be investigated. Heterospectral correlation is a very promising development in the analysis of protein vibrational spectra whenever a protein system under the same perturbations is analyzed for multiple spectroscopic probes characteristic for two different types of closely related spectroscopies. McClure et al. tested the initial application of the hetero 2DCoS to a protein [134]. They utilized the concept of heterospectral correlation for near-IR ↔ Raman spectra of mixtures of lysozyme and sugar (99 % sucrose) in the solid state to improve the interpretation of the NIR spectra of the two compounds. The most probable origin of the overtones

and combination bands from lysozyme and sucrose has been examined by connecting the NIR spectra, which are composed of rather broad and overlapping bands, to the Raman spectra, which are composed of sharp, clearly separated and well-assigned bands. The analysis has been restricted to statistical 2D correlation and to slice spectra extracted at specific frequencies of sucrose or lysozyme. Generally, the slice spectrum is obtained by selecting the frequency of a well-defined vibrational mode in one region whereas the frequencies in the other region change. The spectrum contains a large amount of information about the relationship between one selected absorber in a region and all of the absorbers in another region. Slice spectra extracted both from the synchronous and asynchronous spectra could be very useful in the interpretation of hetero 2DCoS results.

For the first time, the asynchronous foundations have been applied in an associated analysis of the near-IR ↔ mid-IR spectra of RNase A [50]. The lack or presence of asynchronous heterocorrelation cross peaks between the mid-IR amide I components assigned to specific secondary structures and the N–H combination bands in the near-IR range determined whether the near-IR bands could have resulted from the same secondary structure.

The heterospectral correlations between the mid-IR ↔ Raman [40, 41], ECD ↔ Raman [42], and mid-IR ↔ VCD [43] spectra of the protein system have been very attractive from the perspective of better understanding the correlation between the bands attributed to the same secondary structure elements in the two types of spectroscopy. Heterospectral correlation has greatly facilitated band assignments, particularly in situations in which a band was well identified in one range/spectroscopy because of a high structural sensitivity but its counterpart could not be easily disclosed in the second range or by the second spectroscopy because of lower optical activity, caused by differences in the selection rules [42, 43]. The mid-IR ↔ Raman 2DCoS has provided new insight into the correlation between the IR and Raman bands in the amide III region and has confirmed the band assignments for the unordered form [40]

Raman and ROA, two spectroscopies that are extremely sensitive to changes in secondary structure, were pooled together to analyze the pH-induced α-helix-to-disordered transition in poly(L-glutamic acid) (PLG) [105] and the temperature-induced α-helix-to-β-sheet transition in poly(L-lysine) [108]. This combination enabled the complete monitoring of structural changes and of spectral changes in a much wider range of bands (from 1,800 to 800 cm^{-1}) than has been practiced in the standard analysis based only on one spectroscopy without the support of a second one. The associated analysis presented the great potential for the discovery of new transition mechanisms in relatively complex biological systems. It has been shown that the pH-induced conformational change in PLG might not be a simple two-state transition from the α-helix to the disordered structure. The apparent inverse order effect was observed during the temperature-induced transition of PLL from an α-helix to a β-sheet. The changes in the β-sheet structure apparently occurred before those in the α-helix. This phenomenon has been the subject of detailed studies performed with the support of a new development in 2DCoS, i.e., kν-correlation analysis [31, 32].

Heterospectral correlation spectroscopy has also been used to establish the relationship between the information extracted from NMR and Raman spectroscopy [104]. In this approach, the structural evolution of silk fibroin as the pH changed from 6.8 to 4.8 was examined by two probes with different frequency scales ($10^{5–9}$ Hz for nuclear spin and $10^{12–14}$ Hz for molecular vibrations). With the two spectroscopies, the changes that develop at different rates at different molecular levels were probed. Three different stages of the transition detected in the 2D NMR homo-spectral correlation have been correlated with specific secondary structure changes from the soluble helix state to the insoluble β-sheet state.

The concept of heterocorrelation has been expanded further to synchronous sample–sample hybrid 2D correlation analysis [135]. This analysis concerns (1) two systems under two different types of perturbations, (2) the same system under related perturbations, (3) two systems under the same perturbation applied under different conditions [76], or (4) two systems under the same perturbation [30]. This hetero variant of 2DCoS is particularly useful for analyses aimed at the comparison of responses to temperature, pressure, H/D exchange, or chemical denaturants by mutants and genetic variants, by the holo- and apo-forms, or by proteins from two different families. Hetero 2DCoS in the different sample/perturbation configuration opens a wide range of correlations, which, compared with the conventional homo 2DCoS, could help give additional insights into the studied systems.

8.4.2 Moving-Window 2D Correlation

A very promising variant form of 2DCoS in protein research is the moving-window 2D correlation introduced by Richardson et al. [136] for the analysis of transitions in liquid crystal systems. Frequently, the interpretation of 2D correlation spectra based on a large set of spectral data becomes difficult. This difficulty could arise if, at different stages of the perturbation, unequal progress of many processes in the studied protein system influence the spectral changes. The 2D analysis of these processes could be rationalized by the subdivision of the full data set into smaller sets; next, these smaller sets can be separately analyzed with the synchronous component of 2DCoS. Each window provides local information about the properties of the protein system. To obtain a global picture of the changes from the subsets of the data, the autopower spectra of the subsets are plotted in a 2D layout. The moving-window 2D correlation approach is more effective when a large data set, covering a whole perturbation, is segmented into a large number of small subsets chosen in a way that ensures the presence of at least two measured spectra in each subset. If the subsets include a large number of spectra, details of the transitional stages may be lost. The application of moving windows aids in the observation of small, intermediate, structural changes, which could not be detected in techniques that only apply to the initial and final results of the conformational transitions in proteins. This approach also allows the spectra from the entire perturbation range to be divided into groups that contain more spectra specific to the intermediates.

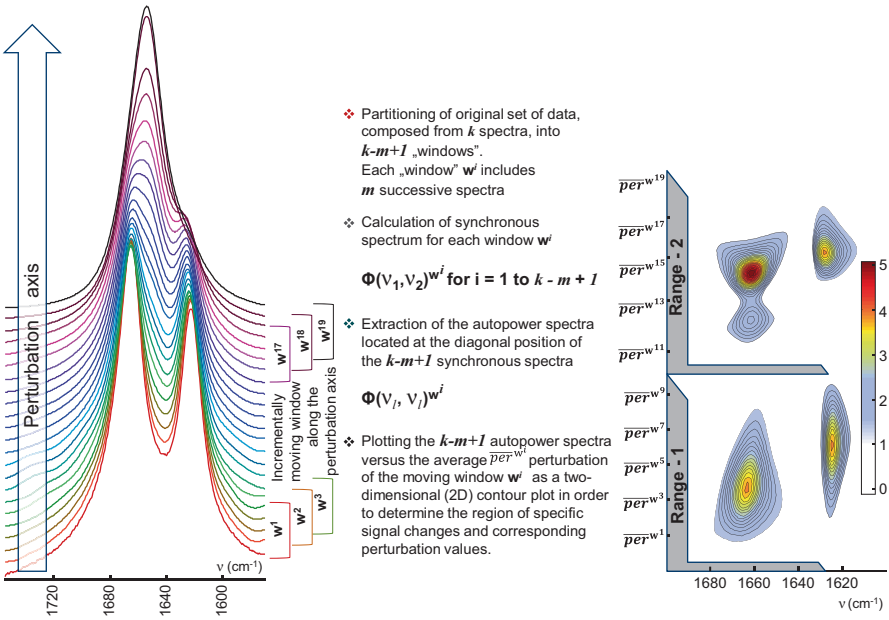


Fig. 8.8 Scheme of the principle of moving-window 2D correlation analysis

The idea behind the procedure is presented in Fig. 8.8 for an example of a sample measured at 23 perturbation steps and whose spectra are sequentially divided into 19 windows containing 5 spectra.

Ashton and Blanch have utilized this 2D correlation approach to detect conformational transitions in poly(L-glutamic acid) and poly(L-lysine) measured as a function of temperature and pH, respectively. It has been possible to determine the temperature and pH values at which the maximum intensity changes occurred and that corresponded to the temperature and pH transition points [110]. Moving-window 2DCoS has also been applied to monitor the pH-induced changes in PLG. This approach disclosed two distinct phases during the unfolding of the helix in the α-helix-to-disordered transition, which occurred at ~pH 4.9 and ~pH 5.2 and may have resulted from the difference in helical stability between the end and central regions of the α-helix [105].

A small protein, α-lactalbumin, has been extensively studied by numerous techniques because of its tendency to form a molten globule under specific conditions. In α-la, the pH-induced conformational transitions from the native-to-molten globule states were investigated by Raman spectra subjected to two-dimensional correlation analysis combined with moving-window correlation [106]. This analysis has identified three distinct transitions occurring at the pH ranges ~pH 6.5–4.6, ~pH 4.6–3.6 and ~pH 3.6–1.8 and has suggested the possible existence of three intermediate conformations. Next, from the synchronous spectra calculated for each range, the autopower spectra have been extracted. The comparative analysis of the autopower

spectra has allowed the postulate of significant structural differences between the first two phases and the last. During the first two transition stages, changes in the secondary structure of the backbone in the β-domain were mainly detected whereas in the final transition, the phase changes also take place in the α-domain and side chain residues dominated by solvent exposure.

8.4.3 Model-based correlation—βν and kν correlation analyses

To overcome some difficulties in the interpretation of temporal or phase relationships between intensity variations within a set of spectra subjected to the generalized 2D approach, Dluhy's group proposed a method in which the experimental intensities are correlated with a reference set of intensities changing according to a sinusoidal [137] (βν-correlation) or exponential function [138] (kν-correlation). Additionally, the global correlation phase angle method has been presented by Dluhy's group as a tool to improve the interpretation of the temporal relationships between the intensity variations [35].

The βν and kν-correlations fulfill the conditions of heterocorrelation; one set is the experimental spectra, and the second set is the basis set of sinusoidal reference functions with systematically varying phase angle (β) or exponential functions with rate constants (k) that undergo systematic changes in a user-defined range. In the proposed method, only the asynchronous correlation spectra are employed; these spectra are generally more sensitive to differences in the form of the signal variation than the synchronous spectra. The effective phase angle, β_e, and the effective rate constant, k_{eff}, determined from the $\Psi(\beta,\nu)$ and $\Psi(k,\nu)$ plots as the points of maximum correlation intensity have been used to verify the sequence of intensity changes, as shown in Fig. 8.9 The spectra subjected to the kν-correlation analysis are superposition of three bands, which order the spectral intensity changes give numbers in Fig. 8.9a. To correctly determine the order of the intensity changes from the asynchronous $\Psi(\nu_1,\nu_2)$, the synchronous sign around the asynchronous peaks must be considered. The grey outlines in Fig. 8.9c mark the areas where the sign was negative. Neglecting this fact results in a false sequence of intensity changes. The application of the kν-correlation procedure in which the spectra (Fig. 8.9a) have been correlated with the exponential runs $\exp(-k\,t + R)$ with different k values (Fig. 8.9b) has allowed the detection of three peaks. The k_{eff} values of these peaks reflect the a priori known sequence of intensity changes (Fig. 8.9d) because events at frequencies with a larger k_{eff} value occur earlier than events at frequencies with smaller k_{eff} value [68].

Because the two correlations are robust and are generally applicable, they have been used in a wide range of studies. On one hand, the generalized 2DCoS is a model-free procedure and directly extracts information from the synchronous and asynchronous spectra. On the other hand, users frequently refer to an assumed model of signal changes, and the 2DCoS pattern of the model is compared with that obtained from the experimental data [37, 39, 63, 71]. For protein spectra, it could be

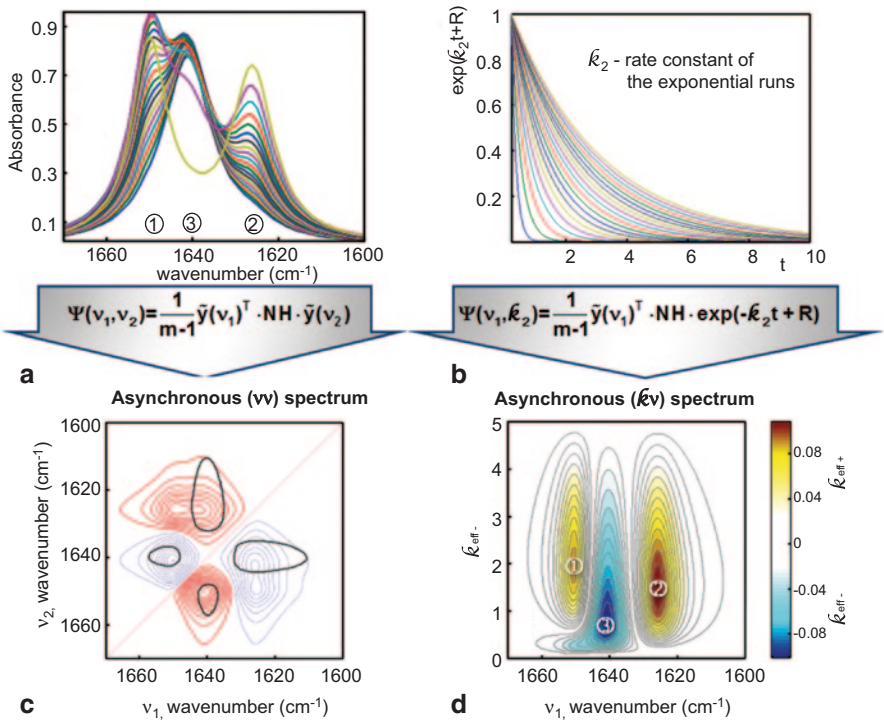


Fig. 8.9 The sequence of events from analysis of the asynchronous spectrum and comparison of the k_{eff} values in the kν-correlation plot. **a** The simulated absorbance spectra with numbers according to the a priori known order, **b** the calculated exponential function for different k values, **c** the asynchronous spectrum with outlines marking the areas where the sign of the synchronous spectrum was negative, **d** the kν-correlation plot with kν-peaks marked based on the k_{eff} values

very difficult task to build a model that is well-matched to the experiment. A priori information about the studied protein system in a form of the relative values of k_{eff} or β_e for all the experimentally detectable structural components of this protein could considerably facilitate the model-building procedure.

The βν-correlation procedure has found an excellent application in monitoring the ordering behavior of phospholipid molecules in bilayers [139, 140] and structural changes in proteins incorporated into an ordered medium. The secondary structure of two small hydrophobic proteins (pulmonary surfactant protein SP-B and SP-C) fused into phospholipid bilayers perturbed by increasing surface pressure was analyzed by βν-correlation [58] and kν-correlation [68]. The Ψ(β,ν) spectrum of the protein showed the presence of two α-helix components—a hydrophobic fraction associated with the lipid chains and a hydrophilic fraction parallel to the membrane surface. The distribution of the Ψ(β,ν) intensity between the two α-helix cross peaks indicated that at low surface pressures, the more hydrophobic helix fraction dominated, whereas the more hydrophilic helix fraction dominated at high surface pressures. 2D kν-correlation analysis has proven that if the studied protein is transferred

into the interior of the lipid bilayers, the α-helix and α-sheet of the protein reorient independently toward the surface normal as the surface pressure increases.

As the next example shows, βν-correlation is very well suited to the analysis of the temporal order of events during temperature-induced changes in the local molecular environment of the protein. The subject of interest was a small designed protein, utilized in the preparation of a "smart" protein micellar system, and because of a reversible thermal transition, this protein triggers rapid changes in the properties of the micelle [72]. The structural units of this protein have been classified according to the β_e values. Analysis of the changes of the units has shown that a thermally induced helix-to-sheet protein folding transition plays the role of a switch for the micelle compacticity and size. The usefulness of the βν- and kν-correlation analyses was also shown through studies of the surface pressure–induced effects on lipid–antibiotic interactions [141] and the influence of temperature on phase transitions in Langmuir polymer films [139].

The kν-correlation analysis has allowed the spectral intensity changes to be quantitatively characterized and has shown excellent performance in the detection of small differences in the rates of dynamic intensity changes in the early stages of fibrillation of lysosome [31, 32, 114]. The Ψ(k,ν) map helped explain the inverse order of events of lysozyme fibrillation deduced from conventional 2DCoS, in which it has been observed that the product of the reaction formed before the initial components of the studied system started to react. This phenomenon, termed apparent inverse order of secondary structural changes [31], was initially described in [61, 103, 108], but at that time, the counterintuitive fact that the appearance of ν-structure occurred before the unfolding of the α-helix seemed to be specific to 2D results whenever aggregation preceded conformational changes within the individual protein molecules. Through application of the kν-correlation analysis, Lednev's group performed a quantitative assessment of the sequence of events using the effective rate constants k_{eff} for each evolution of the secondary structure. The effective rate constants were determined from the best match of the 2D and kν correlation maps for the experimental and model-specific simulated data. Based on the k_{eff} values, it was clear that "Apparent Inverse Order of Events is the Intrinsic Property of Multistep Chemical Reactions" [31] and is typical of all chemical reactions that involve the formation of intermediate(s); this effect may lead to the serious misinterpretation of 2D correlation results. The proposed method of analysis is not limited to a particular spectroscopic technique and can be used to elucidate the mechanisms of complex chemical processes. The software used to justify the inverse sequential order effect is available as supporting material to [31].

8.4.4 Chemometric Procedures

A very useful development in 2D correlation spectroscopy in protein research has been obtained by the incorporation of multivariate chemometric techniques, especially principal component analysis (PCA). Moreover, as shown by the results ob-

tained for other types of systems [142–145] or for different spectroscopies [102], multivariate curve resolution (MCR), mainly of pure concentration profiles, could also contribute valuable information useful in the interpretation of infrared 2DCoS results. It can be expected that the 2DCoS-MCR combination will find greater applications in the analysis of infrared 2DCoS studies of protein, as in [70]. Numerous literature data from Tauler's group and our results have shown that MCR itself has great potential in protein investigations [146, 147].

Despite their complementarity, 2DCoS and PCA are totally independent computational methods. In PCA, as in many other chemometric methods, each spectrum in the entire range is considered to be one integral object represented by scores and loading vectors. In contrast, 2DCoS is not based on the integral approach, and the spectral variations at each frequency contribute individually to the synchronous and asynchronous values. This combination of 2DCoS with PCA or another chemometric method is highly recommended because it supports the reliability of obtained results and primarily aids in the interpretation of the sequence of spectral intensity changes, which are determined from the so-called Noda's rules. Moreover, a plot of the loading values as a function of the original variables (frequencies) helps to deconvolute the amide bands and assign the bands to successive perturbation steps. Comparative analysis of the characteristics of PCA and 2DCoS has been performed for the near-IR spectra of HSA measured as a function of concentration at room temperature [92]. The PCA model, which is based on the two component provided matrix of two loadings plotted as a function of frequency, has been compared with the autopower spectrum and slice spectra extracted from the asynchronous spectrum at the frequency corresponding to the maximum asynchronicity. A strong similarity has been observed between the autopower spectrum and the first PCA loading vector. Asynchronous changes from the slice spectrum were well matched with the second PCA loading vector, which is orthonormal to the first vector. This analysis is supported by the distribution of the score values, which shows the relationships between the samples and provides additional information that could be very helpful in the interpretation of the sequence of spectral intensity changes from the 2DCoS results. In most cases, if the two methods were combined, only the scores plot, which shows a possible cluster of samples in response to the perturbation, has been analyzed. 2DCoS of spectra separated into clusters has provided specific information that is strictly correlated with a well-defined perturbation range [30, 57, 74, 76, 77, 83, 87, 98]. For 2DCoS of data covering the entire perturbation, some information could be unnoticed, mainly information from intermediate states with low population that occur in a narrow perturbation range. PCA has also been employed at the data pretreatment level to remove noise from spectra before 2DCoS [148].

An overview of the different modern statistical and numerical methods used in the analysis of spectral data and, in particular, for the quantitative characterization of protein structural evolution is given in [149]. This survey provides valuable guidelines for biospectroscopists who would like to use these powerful methods in their studies. Such information can also be found in [150], which provides comprehensive information on procedures oriented toward improving the interpretation

and quantitative prediction of secondary structure from infrared spectra. Additionally, the strengths and weaknesses of these procedures are discussed.

Acknowledgments This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2012R1A1A3011844) and the Korean Federation of Science and Technology Societies (KOFST) grant funded by Korea Government (MEST, Basic Research Promotion Fund). BC-M is grateful to the National Science Centre for research funding (grant no. 2012/05/B/ST/02029).

References

1. Noda I, Ozaki Y (2004) Two-dimensional correlation spectroscopy: applications in vibrational and optical spectroscopy. John Wiley & Sons, Chichester
2. Baiz CR, Reppert M, Tokmakoff A (2013) An introduction to protein 2d ir spectroscopy. In: Fayer MD (ed) Ultrafast infrared vibrational spectroscopy, CRC Press. http://web.mit.edu/~tokmakofflab/documents/Intro2DIR.pdf. Accessed 10 Apr 2013
3. Yang WJ, Griffiths P, Byler DM, Susi H (1885) Protein conformation by infrared spectroscopy: resolution enhancement by fourier self-deconvolution. Appl Spectrosc 39(2):282–287
4. Jackson M, Mantsch HH (1995) The use and misuse of FTIR spectroscopy in the determination of protein structure. Crit Rev Biochem Mol Biol 30(2):95–120
5. Noda I (1986) Two-dimensional infrared (2D IR) spectroscopy of synthetic and biopolymers. Bull Am Phys Soc 31(3):520
6. Noda I (1993) Generalized two-dimensional correlation method applicable to infrared, Raman, and other types of spectroscopy. Appl Spectrosc 47(9):1329–1336
7. Raval A, Piana S, Eastwood MP, Dror RO, Shaw DE (2012) Refinement of protein structure homology models via long, all-atom molecular dynamics simulations. Proteins 80(8):2071–2079
8. Dobson CM (2003) Protein folding and misfolding. Nature 426(6968):884–90
9. Dill KA, MacCallum JL (2012) The Protein-Folding Problem, 50 Years On. Science 338(6110):1042–1046
10. Xu M, Ermolenkov VV, Uversky VN, Lednev IK (2008). Hen egg white lysozyme fibrillation: a deep UV resonance Raman spectroscopic study. J Biophotonics 1(3):215–229
11. Cerf E, Sarroukh R, Tamamizu-Kato S, Breydo L, Derclaye S, Dufrêne YF, Narayanaswami V, Goormaghtigh E, Ruysschaert JM, Raussens V (2009) Antiparallel β-sheet: a signature structure of the oligomeric amyloid β-peptide. Biochem J 421(3):415–423
12. Zandomeneghi G, Krebs MR, McCammon MG, Fandrich M (2004) FTIR Reveals Structural Differences between Native β-Sheet Proteins and Amyloid Fibrils. Protein Sci 13(12):3314–3321
13. Meier RJ (2005) Vibrational spectroscopy: a 'vanishing' discipline? Chem Soc Rev 34(9):743–752
14. Noda I (2002) General Theory of Two-dimensional (2D) Analysis. In: Chalmers JM, Griffiths PR (eds) Handbook of Vibrational Spectroscopy, vol. 3. John Wiley & Sons, p 2123–2134
15. Ozaki Y (2002) 2D Correlation spectroscopy in vibrational spectroscopy. In: Chalmers JM, Griffiths PR (eds) Handbook of vibrational spectroscopy, vol. 3. John Wiley & Sons, pp 2135–2172
16. Noda I (2008) Generalized two-dimensional correlation spectroscopy. In: Laane J (ed) Frontiers of molecular spectroscopy. Elsevier Science, p 367–381
17. Noda I (2004) Advances in two-dimensional correlation spectroscopy. Vib Spectrosc 36(2):143–165

18. Noda I (2006) Progress in two-dimensional (2D) correlation spectroscopy. J Mol Struct 799(1–3):2–15
19. Noda I (2008) Recent advancement in the field of two-dimensional correlation spectroscopy. J Mol Struct 883–884:2–26
20. Noda I (2010) Two-dimensional correlation spectroscopy-Biannual survey 2007-2009. J Mol Struct 974(1–3):3–24
21. Noda I (2012) Close-up view on the inner workings of two-dimensional correlation spectroscopy. Vib Spectrosc 60:146–153
22. Ozaki Y, Šašic S, Tanaka T, Noda I (2001) Two-Dimensional correlation spectroscopy: principle and recent theoretical development. Bull Chem Soc Jpn 74(1):1–17
23. Xiaoming D, Bo Y, Haiying Z, Guangzhong Y, Ozaki Y (2004) Generalized two–dimensional correlation spectroscopy—theory and applications in analytical field. Sci China, Ser B: Chem 47(3):257–266
24. Czarnecki MA (2011) Two-dimensional correlation analysis of hydrogen-bonded systems: basic molecules. Appl Spectrosc Rev 46(1):67–103
25. Pazderka T, Kopecký V Jr (2008) 2D correlation spectroscopy and its application in vibrational spectroscopy using MATLAB. http://dsp.vscht.cz/konference_matlab/MATLAB08/prispevky/081_pazderka.pdf. Accessed 25 Mar 2013
26. Czarnik-Matusewicz, Pilorz S, Ashton L, Blanch EW (2006) Potential pitfalls concerning visualization of the 2D results. J Mol Struct 799(1–3):253–258
27. Huang H (2007) "Sequential Order" rules in generalized two-dimensional correlation spectroscopy. Anal Chem 79(21):8281–8292
28. Ma L, Ahmed Z, Asher SA (2011) Ultraviolet resonance raman study of side chain electrostatic control of poly-l-lysine conformation. J Phys Chem B 115(14):4251–4258
29. Huang H, Ding X, Zhu C, He Z, Yu Y (2013) 2D correlation analysis: sequential order judging. Anal Chem 85(4):2161–2168
30. Czarnik-Matusewicz B, Kim SB, Jung YM (2009) A study of urea-dependent denaturation of β-lactoglobulin by principal component analysis and two-dimensional correlation spectroscopy. J Phys Chem B 113(2):559–566
31. Shashilov V, Lednev IK (2008) 2D Correlation deep uv resonance raman spectroscopy of early events of lysozyme fibrillation: kinetic mechanism and potential interpretation pitfalls. J Am Chem Soc 130(1):309–317
32. Shashilov VA, Lednev IK (2009) Two-dimensional correlation Raman spectroscopy for characterizing protein structure and dynamics. J Raman Spectrosc 40(12):1749–1758
33. Ma L, Sikirzhytski V, Hong Z, Lednev IK, Asher SA (2013) Insight into resolution enhancement in generalized two dimensional correlation spectroscopy. Appl Spectrosc 67(3):283–290
34. Jung YM, Noda I (2006) New approaches to generalized two-dimensional correlation spectroscopy and its applications. Appl Spectrosc Rev 41(5):515–547
35. Dluhy R, Shanmukh S, Morita SI (2006) The application of two-dimensional correlation spectroscopy to surface and interfacial analysis. Surf Interface Anal 38(11):1481–1496
36. Nabet A, Pézolet M (1997) Two-dimensional FT-IR spectroscopy: a powerful method to study the secondary structure of proteins using H-D exchange. Appl Spectrosc 51(4):466–469
37. Sefara NL, Magtoto NP, Richardson HH (1997) Structural characterization of beta-lactoglobulin in solution using two-dimensional ft mid-infrared and ft near-infrared correlation spectroscopy. Appl Spectrosc 51(4):536–540
38. Schultz CP, Fabian H, Mantsch HH, (1998) Two-dimensional mid-IR and near-IR correlation spectra of ribonuclease a: using overtones and combination modes to monitor changes in secondary structure. Biospectroscopy 4:S19–S29
39. Fabian H, Mantsch HH, Schultz CP (1999) Two-dimensional IR correlation spectroscopy: sequential events in the unfolding process of the λ Cro-V55C repressor protein. Proc Natl Acad Sci USA 96(23):13153–13158

40. Jung Y, Czarnik-Matusewicz B, Ozaki Y (1999) An IR vs. Raman 2D heterospectral correlation study of the secondary structure of β-lactoglobulin in aqueous solutions. In: Ozaki Y, Noda I (eds) AIP Conference Proceedings 503. Two-Dimensional Correlation Spectroscopy. Kobe-Sanda, pp 275-278 (29 August–1 September 1999)
41. Pancoska P, Kubelka J, Keiderling TA (1999) Novel use of a static modification of two-dimensional correlation analysis. Part I: comparison of the secondary structure sensitivity of electronic circular dichroism, ft-ir, and raman spectra of proteins. Appl Spectrosc 53(6):655–665
42. Kubelka J, Pancoska P, Keiderling TA (1999) Novel use of a static modification of two-dimensional correlation analysis. Part II: Hetero-spectral correlations of protein raman, ft-ir, and circular dichroism spectra. Appl Spectrosc 53(6):666–671
43. Kubelka J, Pancoska P, Keiderling TA (1999) In: Greve J, Puppels GJ, Otto C (eds) Spectroscopy of biological molecules: new directions. 8th European Conference on the Spectroscopy of Biological Molecules, Enschede 29 August–2 September 1999. Kluwer Academic Publishers, pp 67–68
44. Meskers S, Ruysschaert JM, Goormaghtigh E (1999) Hydrogen-deuterium exchange of streptavidin and its complex with biotin studied by 2D-Attenuated total reflection Fourier transform infrared spectroscopy. J Am Chem Soc 121(22):5115–5122
45. Smeller L, Heremans K (1999) 2D FT-IR spectroscopy analysis of the pressure-induced changes in proteins. Vib Spectrosc 19(2):375–378
46. Smeller L, Heremans K (1999) Hydrogen-deuterium exchange versus conformational changes in proteins. A two-dimensional ftir approach. In: H. Ludwig (ed) Advances in high pressure bioscience and biotechnology. Springer-Verlag, Berlin, pp 223–226
47. Czarnik-Matusewicz B, Murayama K, Wu Y, Ozaki Y (2000) Two-dimensional attenuated total reflection/infrared correlation spectroscopy of adsorption-induced and concentration-dependent spectral variations of β-Lactoglobulin in aqueous solutions. J Phys Chem B 104(32):7803–7811
48. Dzwolak W, Kato M, Shimizu A, Taniguchi Y (2000) Comparative two-dimensional Fourier transform infrared correlation spectroscopic study on the spontaneous, pressure-, and temperature-enhanced H/D exchange in alpha-lactalbumin. Appl. Spectrosc 54(7):963–967
49. Jung YM, Czarnik-Matusewicz B, Ozaki Y (2000) Two-dimensional infrared, two-dimensional raman, and two-dimensional infrared and raman heterospectral correlation studies of secondary structure of β-lactoglobulin in buffer solutions. J Phys Chem B 104(32):7812–7817
50. Schultz CP, Bârzu O, Mantsch HH (2000) Two-dimensional infrared correlation analysis of protein unfolding: use of spectral simulations to validate structural changes during thermal denaturation of bacterial CMP kinases. Appl Spectrosc 54(7):931–938
51. Smeller L, Rubens P, Frank J, Fidy J, Heremans K (2000) Two dimensional Fourier-transform infrared correlation spectroscopy study of the high-pressure tuning of proteins. Vib Spectrosc 22:119–125
52. Sonoyama M, Nakano T (2000) Infrared rheo-optics of bombyx mori fibroin film by dynamic step-scan FT-IR spectroscopy combined with digital signal processing. Appl Spectrosc 54(7):968–973
53. Filosa A, Wang Y, Ismail AA, English AM (2001) Two-dimensional infrared correlation spectroscopy as a probe of sequential events in the thermal unfolding of cytochromes c. Biochemistry 40(28):8256–8263
54. Lefèvre T, Subirade M (2001) Conformational rearrangement of β-lactoglobulin upon interaction with an anionic membrane. Biochim Biophys Acta 1549(1):37–50
55. Murayama K, Wu Y, Czarnik-Matusewicz B, Ozaki Y (2001) Two-dimensional/attenuated total reflection infrared correlation spectroscopy studies on secondary structural changes in human serum albumin in aqueous solutions: pH-dependent structural changes in the secondary structures and in the hydrogen bondings of side chains. J Phys Chem B 105(20):4763–4769

56. Paquet MJ, Laviolette M, Pézolet M, Auger M (2001) Two-dimensional infrared correlation spectroscopy study of the aggregation of cytochrome c in the presence of dimyristoylphosphatidylglycerol. Biophys J 81(1):305–312
57. Wu Y, Murayama K, Ozaki Y (2001) Two-dimensional infrared spectroscopy and principle component analysis studies of the secondary structure and kinetics of hydrogen–deuterium exchange of human serum albumin. J Phys Chem B 105(26):6251–6259
58. Shanmukh S, Howell P, Baatz JE, Dluhy RA (2002) Effect of hydrophobic surfactant proteins SP-B and SP-C on phospholipid monolayers. Protein structure studied using 2D IR and βν correlation analysis. Biophys J 83(4):2126–2141
59. Turnay J, Olmo N, Gasset M, Iloro I, Arrondo JL, Lizarbe MA (2002) Calcium-dependent conformational rearrangements and protein stability in chicken annexin. Biophys J 83(4):2280–2291
60. Wu Y, Murayama K, Czarnik-Matusewicz B, Ozaki Y (2002) Two-dimensional attenuated total reflection/infrared correlation spectroscopy studies on concentration and heat-induced structural changes of human serum albumin in aqueous solutions. Appl Spectrosc 56(9):1186–1193
61. Yan YB, Wang Q, He HW, Hu XY, Zhang RQ, Zhou HM (2003) Two-dimensional infrared correlation spectroscopy study of sequential events in the heat-induced unfolding and aggregation process of myoglobin. Biophys J 85(3):1959–1967
62. Iloro I, Chehin R, Goni FM, Pajares MA, Arrondo JL (2004) Methionine adenosyltransferase α-helix structure unfolds at lower temperatures than β-Sheet: a 2D-IR Study. Biophys J 86(6):3951–3958
63. Lefèvre T, Arseneault K, Pézolet M (2004) Study of protein aggregation using two-dimensional correlation infrared spectroscopy and spectral simulations. Biopolymers 73(6):705–715
64. Yan YB, Wang Q, He HW, Zhou HM (2004) Protein thermal aggregation involves distinct regions: sequential events in the heat-induced unfolding and aggregation of hemoglobin. Biophys J 86(3):1682–1690
65. Iloro I, Goni FM, Arrondo JL (2005) A 2D-IR study of heat- and [^{13}C] urea-induced denaturation of sarcoplasmic reticulum Ca2+-ATPase. Acta Biochim Pol 52(2):477–483
66. Ortiz M, Sanoguet Z, Hu H, Chazin WJ, McMurray CT, Salisbury JL, Pastrana-Rios B (2005) Dynamics of hydrogen–deuterium exchange in chlamydomonas centrin. Biochemistry 44(7):2409–2418
67. Richard JA, Kelly I, Marion D, Auger M, Pézolet M (2005) Structure of β-purothionin in membranes: a two-dimensional infrared correlation spectroscopy study. Biochemistry 44(1):52–61
68. Shanmukh S, Biswas N, Waring AJ, Walther FJ, Wang Z, Chang Y, Notter RH, Dluhy RA (2005) Structure and properties of phospholipid-peptide monolayers containing monomeric SP-B1-25: II. Peptide conformation by infrared spectroscopy. Biophys Chem 113(1):233–244
69. Zhang J, Yan YB (2005) Probing conformational changes of proteins by quantitative second-derivative infrared spectroscopy. Anal Biochem 340(1):89–98
70. Domínguez-Vidal A, Saenz-Navajas MP, Ayora-Cañada MJ, Lendl B (2006) Detection of albumin unfolding preceding proteolysis using fourier transform infrared spectroscopy and chemometric data analysis. Anal Chem 78(10):3257–3264
71. Iloroa I, Pastrana-Rios B (2006) Simulation of FT-IR spectra and 2D-COS analysis for the H/D exchange of two related ligands. J Mol Struct 799(1–3):153–157
72. Sallach RE, Wei M, Biswas N, Conticello VP, Lecommandoux S, Dluhy RA, Chaikof EL (2006) Micelle density regulated by a reversible switch of protein. J Am Chem Soc 128(36):12014–12019
73. Sánchez-Bautista S, Kazaks A, Beaulande M, Torrecillas A, Corbalán-García S, Gómez-Fernández JC (2006) Structural study of the catalytic domain of PKCzeta using infrared spectroscopy and two-dimensional infrared correlation spectroscopy. FEBS J 273(14):3273–86

74. Yan YB, Zhang J, He HW, Zhou HM (2006) Oligomerization and aggregation of bovine pancreatic ribonuclease a: characteristic events observed by FTIR spectroscopy. Biophys J 90(7):2525–2533
75. Bernabeua A, Contrerasb LM, Villalaín J (2007) Two-dimensional infrared correlation spectroscopy study of the interaction of oxidized and reduced cytochrome c with phospholipid model membranes. Biochim Biophys Acta, Biomembr 1768(10):2409–2420
76. Kamerzell TJ, Middaugh CR (2007) Two-dimensional correlation spectroscopy reveals coupled immunoglobulin regions of differential flexibility that influence stability. Biochemistry 46(34):9762–9773
77. Rodríguez-Casado A, Molina M, Carmona P (2007) Spectroscopic study of conformational changes accompanying self-assembly of hcv core protein. Proteins: Struct, Funct, Bioinf 66(1):110–117
78. Iloro I, Narváez D, Guillén N, Camacho CM, Guillén L, Cora E, Pastrana-Ríos B (2008) The kinetics of the hydrogen/deuterium exchange of epidermal growth factor receptor ligands. Biophys J 94(10):4041–4055
79. Massaro S, Zlateva T, Torre V, Quaroni L (2008) Detection of molecular processes in the intact retina by ATR-FTIR spectromicroscopy. Anal Bioanal Chem 390(1):317–322
80. Arada I, Julien JP, Torre BG, Huarte N, Andreu D, Pai EF, Arrondo JLR, Nieva JL (2009) Structural constraints imposed by the conserved fusion peptide on the HIV-1 gp41 epitope recognized by the broadly neutralizing antibody 2F5. J Phys Chem B 113(41):13626–13637
81. Cerdà-Costa N, De la Arada I, Avilés FX, Arrondo JLR, Villegas S (2009) Influence of aggregation propensity and stability on amyloid fibril formation as studied by fourier transform infrared spectroscopy and two-dimensional COS analysis. Biochemistry 48(44):10582–10590
82. Saguer E, Alvarez P, Sedman J, Ramaswamy HS, Ismail AA (2009) Heat-induced gel formation of plasma proteins: new insights by FTIR 2D correlation spectroscopy. Food Hydrocolloids 23(3):874–879
83. Wang LX, Zhang L, Wu Y (2009) Thermally induced early events of ribonuclease a under reducing conditions: evidenced by principal component analysis and two-dimensional correlation. Vib Spectrosc 51(1):4–10
84. Fabian H, Yu Z, Wang Y, Naumanna D (2010) Generalized 2D and time-resolved FTIR studies of protein unfolding events. J Mol Struct 974(1–3):203–209
85. del Valle Sosa L, Alfaro E, Santiago J, Narváez D, Rosado MC, Rodríguez A, Gómez AM, Schreiter ER, Pastrana-Ríos B (2011) The structure, molecular dynamics, and energetics of centrin-melittin complex. Proteins: Struct, Funct, Bioinf 79(11):3132–3143
86. Nowak C, Laredo T, Gebert J, Lipkowski J, Gennis RB, Ferguson-Miller S, Knoll W, Naumann RLC (2011) 2D-SEIRA spectroscopy to highlight conformational changes of the cytochrome c oxidase induced by direct electron transfer. Metallomics 3(6):619–627
87. Zhang M, Zhang L, Wu Y (2011) The pressure tolerance of different poly-l-lysine conformers in aqueous solution: Infrared spectroscopy and two-dimensional correlation analysis. Vib Spectrosc 57(2):319–325
88. Saguer E, Alvarez P, Ismail AA (2012) Heat-induced denaturation/aggregation of porcine plasma and its fractions studied by FTIR spectroscopy. Food Hydrocolloids 27(1):208–219
89. Zhang M, Dang YQ, Liu TY, Li HW, Wu Y, Li Q, Wang K, Zou B (2013) Pressure-induced fluorescence enhancement of the BSA-protected gold nanoclusters and the corresponding conformational changes of protein. J Phys Chem C 117(1):639–647
90. Wang Y, Murayama K, Myojo Y, Tsenkova R, Hayashi N, Ozaki Y (1998) Two-dimensional fourier transform near-infrared spectroscopy study of heat denaturation of ovalbumin in aqueous solutions. J Phys Chem B 102(34):6655–6662
91. Ozaki Y, Murayama K, Wang Y (1999) Application of two-dimensional near-infrared correlation spectroscopy to protein research. Vib Spectrosc 20(2):127–132
92. Murayama K, Czarnik-Matusewicz B, Wu Y, Tsenkova R, Ozaki Y (2000) Comparison between conventional spectral analysis methods, chemometrics, and two-dimensional correlation spectroscopy in the analysis of near-infrared spectra of protein. Appl Spectrosc 54(7):978–985

93. Wu Y, Czarnik-Matusewicz B, Murayama K, Ozaki Y (2000) Two-dimensional near-infrared spectroscopy study of human serum albumin in aqueous solutions: using overtones and combination modes to monitor temperature-dependent changes in the secondary structure. J Phys Chem B 104(24):5840–5847
94. Czarnik-Matusewicz B (2002) Near-infrared spectroscopy combined with two-dimensional correlation spectroscopy in studies of biomolecules. In: Tsuchikawa S (ed) Useful and advanced information in the field of near infrared spectroscopy. Research Signpost, pp 189–216
95. Murayama K, Ozaki Y (2002) Two-dimensional near-IR correlation spectroscopy study of molten globule-like state of ovalbumin in acidic pH region: Simultaneous changes in hydration and secondary structure. Biopolymers 67(6):394–405
96. Yuan B, Murayama K, Wu Y, Tsenkova R, Dou X, Era S (2003) Temperature-dependent near-infrared spectra of bovine serum albumin in aqueous solutions: spectral analysis by principal component analysis and evolving factor analysis. Appl Spectrosc 57(10):1223–1229
97. Yu L, Xiang B (2008) Two-dimensional near-IR correlation spectroscopy study the interaction of protein and famotidine in acidic pH region. Spectrochim Acta A 69(2):599–603
98. Kamerzell TJ, Kanai S, Liu J, Shire SJ, Wang YJ (2009) Increasing IgG concentration modulates the conformational heterogeneity and bonding network that influence solution properties. J Phys Chem B 113(17):6109–6118
99. Mo C, Wu P, Chen X, Shao Z (2009) The effect of water on the conformation transition of Bombyx mori silk fibroin. Vib Spectrosc 51(1):105–109
100. Ma S, Freedman TB, Cao X, Nafie LA (2006) Two-dimensional vibrational circular dichroism correlation spectroscopy: pH-induced spectral changes in L-alanine. J Mol Struct 799(1–3):226–238
101. Ryu SR, Czarnik-Matusewicz B, Dukor RK, Nafie LA, Jung YM (2012) Analysis of the molten globule state of bovine α-lactalbumin by using vibrational circular dichroism. Vib Spectrosc 60:68–72
102. Czarnik-Matusewicz B, Pilorz S (2006) 2DCOS and MCR-ALS as a combined tool of analysis of β-lactoglobulin CD spectra. J Mol Struct 799(1–3):211–220
103. Zhou P, Xie X, Knight DP, Zong XH, Deng F, Yao WH (2004) Effects of pH and calcium ions on the conformational transitions in silk fibroin using 2D Raman correlation spectroscopy and 13C Solid-State NMR. Biochemistry 43(35):11302–11311
104. Hu BW, Zhou P, Noda I, Ruan QX (2006) Generalized two-dimensional correlation analysis of nmr and raman spectra for structural evolution characterizations of silk fibroin. J Phys Chem B 110(36):18046–18051
105. Ashton L, Barron LD, Hecht L, Hyde J, Blanch EW (2007) Two-dimensional Raman and Raman optical activity correlation analysis of the alpha-helix-to-disordered transition in poly(L-glutamic acid). Analyst 132(5):468–79
106. Ashton L, Blanch EW (2010) pH-induced conformational transitions in α-lactalbumin investigated with two-dimensional Raman correlation variance plots and moving windows. J Mol Struct 974(1–3):132–138
107. Ashton L, Johannessen C, Goodacre R (2011) The importance of protonation in the investigation of protein phosphorylation using raman spectroscopy and raman optical activity. Anal Chem 83(20):7978–7983
108. Ashton L, Barron LD, Czarnik-Matusewicz B, Hecht L, Hyde J, Blanch EW (2006) Two-dimensional correlation analysis of Raman optical activity data on the α-helix-to-β-sheet transition in poly(L-lysine). Mol Phys 104(9):1429–1445
109. Ashton L, Czarnik-Matusewicz B, Blanch EW (2006) Application of two-dimensional correlation analysis to Raman optical activity. J Mol Struct 799(1–3):61–71
110. Ashton L, Blanch EW (2008) Investigation of polypeptide conformational transitions with two-dimensional Raman optical activity correlation analysis, applying autocorrelation and moving window approaches. Appl Spectrosc 62(5):469–75
111. Pazderka T, Kopecký V Jr (2010) Two-dimensional correlation analysis of Raman optical activity—Basic rules and data treatment. Vib Spectrosc 60:193–199

112. Shashilov V, Xu M, Ermolenkov VV, Fredriksen L, Lednev IK (2007) Probing a Fibrillation Nucleus Directly by Deep Ultraviolet Raman Spectroscopy. J Am Chem Soc 129(22):6972–6973
113. Lednev IK, Xu M, Shashilov V (2009) Ultraviolet Raman spectroscopy is uniquely suitable for studying amyloid diseases. Curr Sci 97(2):180–185
114. Oladepo SA, Xiong K, Hong Z, Asher SA, Handen J, Lednev IK (2012) UV Resonance Raman Investigations of Peptide and Protein Structure and Dynamics. Chem Rev 112(5):2604–2628
115. Wang G, Geng L (2005) Statistical and Generalized Two-Dimensional Correlation Spectroscopy of Multiple Ionization States. Fluorescence of Neurotransmitter Serotonin. Anal Chem 77(1):20–29
116. Fukuma H, Nakashima K, Ozaki Y, Noda I (2006) Two-dimensional fluorescence correlation spectroscopy IV: Resolution of fluorescence of tryptophan residues in alcohol dehydrogenase and lysozyme. Spectrochim Acta A 65(3–4):517–522
117. Wang G, Gao Y, Geng ML (2006) Generalized two-dimensional heterocorrelation analysis of spectrally resolved and temporally resolved fluorescence of the 8-Anilino-1-naphthalenesulfonate-Apomyoglobin complex with pH perturbation. J Phys Chem B 110(16):8506–8512
118. Chen C, Xiang B, Yu L, Wang T, Zhao B (2008) The application of two-dimensional fluorescence correlation spectroscopy on the interaction between bovine serum albumin and paeonolum in the presence of Fe(III). Spectrosc Lett 41(8):385–392
119. Ozaki Y, Noda I (2000) Two-dimensional vibrational correlation spectroscopy in biomedical sciences. In: Meyers RA (ed) Encyclopedia of analytical chemistry. John Wiley & Sons, pp 322–340
120. Ozaki Y, Murayama K, Wu Y, Czarnik-Matusewicz B (2003) Two-dimensional infrared correlation spectroscopy studies on secondary structures and hydrogen bondings of side chains of proteins. Spectroscopy 17(2–3):79–100
121. Arrondo JLR, Iloro I, Garcia-Pacios M, Goñi FM (2006) Two-dimensional infrared correlation spectroscopy. In: Arrondo JLR, Alonso A (eds). Advanced techniques in biophysics. Springer Series in Biophysics Volume 10, pp 73–88
122. Ozaki Y, Šašic S (2006) Two-dimensional correlation spectroscopy of biological and polymeric materials. In: Braiman M, Gregoriou VG (ed) Vibrational spectroscopy of biological and polymeric materials. CRC Press, Taylor & Francis Group, pp 163–214
123. Vigano C, Manciu L, Buyse F, Goormaghtigh E, Ruysschaert JM (2001) Attenuated total reflection IR spectroscopy as a tool to investigate the structure, orientation and tertiary structure changes in peptides and membrane proteins. Pept Sci 55(5):373–380
124. Goormaghtigh E (2009) FTIR data processing and analysis tools. In: Barth H, Haris PI (eds) Biological and biomedical infrared spectroscopy. Advances in Biomedical Spectroscopy, vol 2. IOS Press, pp 104–128
125. Goormaghtigh E, Raussens V, Ruysschaert JM (1999) Attenuated total refection infrared spectroscopy of proteins and lipids in biological membranes. Biochim Biophys Acta, Rev Biomembr 1422(2):105–185
126. Goormaghtigh E, Cabiaux V, Ruysschaert JM (1994) Determination of soluble and membrane protein structure by Fourier transform infrared spectroscopy. III. Secondary structures. In: Hilderson HJ, Ralston GB (eds) Physicochemical methods in the study of biomembranes. Subcellular Biochemistry, vol 23. Springer, pp 405–450
127. Invernizzi G, Papaleo E, Sabate R, Ventura S (2012) Protein aggregation: Mechanisms and functional consequences. Int J Biochem Cell Biol 44(9):1541–1554
128. Oberg KA, Fink AL (1998) A new attenuated total reflectance Fourier transform infrared spectroscopy method for the study of proteins in solution. Anal Biochem 256(1):92–106
129. Goormaghtigh E, Cabiaux V, Ruysschaert JM (1994) Determination of soluble and membrane protein structure by Fourier transform infrared spectroscopy. II. Experimental aspects, side chain structure, and H/D exchange. In: Hilderson HJ, Ralston GB (eds) Physicoche-

mical methods in the study of biomembranes. Subcellular Biochemistry, vol 23. Springer, pp 363–403
130. Dzwolak W, Kato M, Taniguchi Y (2002) Fourier transform infrared spectroscopy in high-pressure studies on proteins. Biochim Biophys Acta 1595(1–2):131–144
131. Barth A, Zscherp C (2002) What vibrations tell us about proteins. Quaterly Rev Biophys 35(4):369–430
132. Jung YM, Czarnik-Matusewicz B, Kim SB (2004) Characterization of concentration-dependent infrared spectral variations of urea aqueous solutions by principal component analysis and two-dimensional correlation spectroscopy. J Phys Chem 108(34):13008–13014
133. Barton FE, Himmelsbach DS, Duckworth JH, Smith MJ (1992) Two-dimensional vibration spectroscopy: correlation of mid- and near-infrared regions. Appl Spectrosc 46(3):420–429
134. McClure WF, Maeda H, Dong J, Liu Y, Ozaki Y (1996) Two-dimensional correlation of fourier transform near-infrared and fourier transform raman spectra i: mixtures of sugar and protein. Appl Spectrosc 50(4):467–475
135. Wu Y, Jiang J H, Ozaki Y (2002) A new possibility of generalized two-dimensional correlation spectroscopy: hybrid two-dimensional correlation spectroscopy. J Phys Chem A 106(11) 2422–2429
136. Thomas M, Richardson HH (2000) Two-dimensional FT-IR correlation analysis of the phase transitions in a liquid crystal, 4′-n-octyl-4-cyanobiphenyl (8CB). Vib Spectrosc 24(1):137–146
137. Elmore DL, Dluhy RA (2001) βν-Correlation analysis: a modified two-dimensional infrared correlation method for determining relative rates of intensity change. J Phys Chem B 105(45):11377–11386
138. Shanmukh S, Dluhy RA (2004) kν correlation analysis. a quantitative two-dimensional IR correlation method for analysis of rate processes with exponential functions. J Phys Chem A 108(26):5625–5634
139. Elmore DL, Shanmukh S, Dluhy RA (2002) A study of binary phospholipid mixtures at the air-water interface using infrared reflection-absorption spectroscopy and 2D IR βν correlation analysis. J Phys Chem A 106(14):3420–3428
140. Litwińczuk A, Petrus J, Czarnik-Matusewicz B (2012) Orientation of solid supported lipid bilayers examined by βν-correlation analysis. Vib Spectrosc 60:173–179
141. Shanmukh S, Dluhy RA (2004) 2D IR analyses of rate processes in lipid–antibiotic monomolecular films. Vib Spectrosc 36(2):167–177
142. Czarnik-Matusewicz, Pilorz S, Hawranek JP (2005) Temperature-dependent water structural transitions examined by near-IR and mid-IR spectra analyzed by multivariate curve resolution and two-dimensional correlation spectroscopy. Anal Chim Acta 544(1–2):15–25
143. Muik B, Lendl B, Molina-Diaz A, Valcarcel M, Ayora-Cañada MJ (2007) Two-dimensional correlation spectroscopy and multivariate curve resolution for the study of lipid oxidation in edible oils monitored by FTIR and FT-Raman spectroscopy. Anal Chim Acta 593(1):54–67
144. Himmelsbach D, Holser RA (2009) Application of 2D correlation spectroscopy with MCR in the preparation of glycerol polyesters. Vib Spectrosc 51(1):142–145
145. Spegazzini N (2010) Two-dimensional infrared correlation spectroscopy and multivariate curve resolution methods: application to quantitative monitoring of curing process. PhD thesis. http://hdl.handle.net/10803/9043. Accessed 10 Apr 2013
146. Navea S, de Juan A, Tauler R (2003) Modeling temperature-dependent protein structural transitions by combined near-IR and Mid-IR spectroscopies and multivariate curve resolution. Anal Chem 75(20):5592–5601
147. Szyc L, Pilorz S, Czarnik-Matusewicz B (2008) FTIR-ATR investigations of an α-helix to β-sheet conformational transition in poly(l-lysine). J Mol Liq 141(3):155–159
148. Jung JM (2003) Principal component analysis based two-dimensional (PCA-2D) correlation spectroscopy: PCA denoising for 2D correlation spectroscopy. Bull Korean Chem Soc 24(9):1345–1350

149. Shashilov VA, Lednev IK (2010) Advanced statistical and numerical methods for spectroscopic characterization of protein structural evolution. Chem Rev 110(10):5692–5713
150. Hering JA, Haris PI (2009) FTIR spectroscopy for analysis of protein secondary structure. In: Barth A, Haris PI (eds) Biological and biomedical infrared spectroscopy. IOP Press, pp 129–167

Part III
Biological Applications

Chapter 9
Qualitative and Quantitative FT-Raman Analysis of Plants

Hartwig Schulz

Abstract Raman spectroscopy has been found to be a reliable and non-destructive method for rapid discrimination of different plant species or chemotypes if characteristic key bands can be observed in the spectrum. Today, even portable Raman spectrometers are available which only need sample amounts of a few microliters or milligrams. In most cases, measurements can be performed directly on plant tissues as well as on fractions isolated from the plant material by hydro-distillation or solvent extraction. Generally, Raman spectroscopy allows obtaining spectra which may present several characteristic key bands of individual plant components. Very often these bands provide very useful information about the chemical composition, including both primary and secondary metabolites occurring in the investigated samples. Based on such markers, spectroscopic analyses in principle allow to discriminate different species, and even to classify chemotypes among the same species. The ability to rapidly monitor various plant components makes it possible to efficiently select high-quality single plants from wild populations as well as progenies of crossing experiments. Furthermore, Raman spectroscopy can also be used by the processing industry in order to perform fast quality checks of incoming raw materials as well as continuous controlling of various production processes. Beside MS and NMR measurements, Raman spectroscopy represents a very important analytical tool in the field of plant metabolic fingerprinting providing an unbiased, global screening approach to classify samples that change in response to the genetic background, various plant diseases or influences by the environment (e.g. various stress effects).

Keywords Plant breeding • Plant cultivation • Metabolow analysis • Quality control • Raman mapping • Raman imaging

9.1 Introduction

Within the past 20 years Raman spectroscopy has become increasing importance in the field of plant analysis aiming to provide a fast and mostly non-destructive classification of plant tissues but also to get new insights into biophysical and

H. Schulz (✉)
Julius Kühn-Institute, Institute for Ecological Chemistry, Plant Analysis and Stored Product Protection, Federal Research Centre for Cultivated Plants, 14195, Berlin, Germany
e-mail: Hartwig.Schulz@jki.bund.de

M. Baranska (ed.), *Optical Spectroscopy and Computational Methods in Biology and Medicine,* Challenges and Advances in Computational Chemistry and Physics 14,
DOI 10.1007/978-94-007-7832-0_9, © Springer Science+Business Media Dordrecht 2014

biochemical processes and a more detailed knowledge concerning plant diseases at the molecular level. Generally, all polarisable molecules occurring in plant tissues generate Raman spectra. Contrary to infrared spectroscopy, it is possible to analyse fresh plants without performing any sample pre-treatments, because of the low polarizability index of water producing negligible Raman signal intensity.

Surface-enhanced Raman scattering (SERS), first observed in 1974, relies on electronic and chemical interactions between the excitation laser light, the analyte molecule and the SERS substrate (gold or silver nanoparticles). Using this technique, the detection sensitivity can be enhanced up to 6–10 orders of magnitude over conventional Raman spectroscopy [1]. Recent development of reproducible SERS substrates provides new challenges to detect individual plant components even when they occur in lower concentration [2–4]. Usually the production of the metal nanoparticles is performed by reduction of the metal salts with a suitable reducing agent. However, it has been found that certain plants such as alfalfa (*Medicago sativa*), carrots (*Daucus carota*) green tea (*Camellia sinensis*) and red cabbage (*Brassica oleracea* var. *capitata* f. *rubra*) are able to reduce spontaneously metal colloids [4–6]. In the absence of the metal colloids, Raman spectra could not be obtained for most of the plant material for excitations at 514, 633 and 785 nm due to the strong fluorescence of individual plant substances. Generally, the presence of the gold colloids enabled strong SERS spectra, whereas the spectra of silver colloids were weaker and exhibited more noise than those registered for gold. The Raman spectrum of carrot was an exception, showing strong carotenoid signals, which were not visible in presence of gold colloids. Most peaks of the SERS spectra were attributed to adenine dinucleotide (NAD) and other adenine-containing materials, but also some contributions from flavins, chlorophyll and lipids were assigned.

In order to describe the individual gene functions, protein regulation and the production of small molecular weight metabolites, the variation within the biological systems has been studied at the genomic, proteomic and metabolomic levels. The metabolome is defined as the total sum of small molecular weight substances (metabolites) present in an organism which participate in metabolic reactions required for growth, maintenance and normal functions [7]. At present, it is estimated that the number of metabolites detected in the plant kingdom may reach a number up to 200,000. For targeted and non-targeted analysis and quantification as well as metabolic fingerprinting, techniques such as Raman and infrared spectroscopy are valuable analytical tools beside NMR and MS. They provide the opportunity for unbiased high-throughput measurements of the whole system [8] and are comparatively inexpensive. In contrast to the other techniques used for metabolome studies, more or less no sample preparation is necessary which also guarantees that all plant substances are analysed without any discrimination related to less solubility.

9.2 Primary Substances

Primary substances of plant origin (carbohydrates, amino acids, fats, oils and proteins) occur in all cells and are responsible for the individual biochemical processes necessary to produce enough energy, to maintain life, or to stimulate growth of the respective organism. Furthermore, from these basic substances numerous secondary metabolites, exhibiting various healthy properties for human beings, are biosynthesized.

9.2.1 Fatty Oils and Waxes

Vegetable oils obtained from various oil plants such as, coconut, corn, high oleic sunflower, olive oil, peanut, Brazil nut, palm, palm kernel, rapeseed, soybean by mechanical or solvent extraction show characteristic Raman bands in the region of 1,440–1,660 cm^{-1} which are related to C=C stretching modes of *cis*-unsaturated fatty acids as well as CH_2-scissoring modes of saturated fatty acids [9]. In addition characteristic signals occur in the region between 2,990 and 3,020 cm^{-1}. The Raman shifts around 3,010 cm^{-1} were assigned as $\nu_{(=C\text{-}H)}$modes of the olefinic triglycerides which can be used as a reliable indicator of the different degrees of unsaturation [9, 10]. Furthermore, these signals can also be used to accurately predict the individual iodine value of vegetable oils as well as to detect and quantify adulterations in olive oil. Other scattering Raman shifts registered at 1,663 and 1,264 cm^{-1} were also found to correlate with the individual fatty acid profiles determined by gas chromatography.

Vegetable oils present a comparatively weak carbonyl Raman band around 1,750 cm^{-1}. Other signals can be observed in the range between 1,650 and 1,670 cm^{-1} which are assigned to $\nu_{(C=C)}$ vibrations of the unsaturated triglycerides. The fingerprint region between 700 and 1,200 cm^{-1} shows the characteristic C-C skeletal and C-O bond vibrations and provides mainly information about the unsaturated fatty acids [9]. Several Raman studies have been performed to determine oil content, humidity and free fatty acid content in olives [11, 12] or to quantify the amount of sunflower oil adulterant in extra-virgin olive oil from various Mediterranean sites [13]. A new method has been also developed to rapidly predict the iodine value in single rapeseeds [14]. This approach is very useful for breeding research and applied breeding to measure the lipid content and composition of a seedling without to destroy the plant tissue. Thus, the authors were able to distinguish seedlings of a classic rapeseed cultivar from a new genetically modified cultivar. This successful discrimination by means of cluster analysis based on Raman data is mainly related to a lower iodine value of the genetically modified cultivar, as it contains myristic acid which is normally not detectable in the fatty acid profile of rape.

Several studies have been performed applying Raman spectroscopy to discriminate between different edible oils and fats [15–17]. Using linear discriminant analysis (LDA) and canonical variate analysis (CVA) adulterants of edible oils

and the individual botanical origin were identified. A similar approach has been presented by Weng et al. [18] to prove the authenticity of *Camelia oleifera* oil used in Taiwan.

A rapid method to detect adulterations of extra virgin olive oil with cheaper vegetable oils has been also described [19]. Based on various extra virgin oil samples exhibiting different amounts of olive oil adulteration, a suitable calibration equation was developed ($R^2 = 0.997$, SEP = 1.72 %) allowing to detect even low amounts of other added vegetable oils. The potential capability of Raman spectroscopy to identify and quantify monoterpenes isolated from aromatic plants in olive oil was studied by Omar et al. [20]. They mixed olive oil with different amounts of rosemary oil, measured in the range between 150 and 3,000 cm^{-1} and developed calibrations for eucalyptol and camphor. The predicted concentrations of essential oil in the analysed olive oil samples were found to correlate very well with the reference GC data. The authors stress that Raman spectroscopy is therefore a useful screening method for various purposes in the field of food chemistry.

Another approach for discrimination of olive oils from various cheaper vegetable oils (sunflower oil, safflower oil, soybean oil, wheat germ oil, walnut oil, flaxseed oil) has been developed by El-Abassy and al. [21]. Raman measurements were carried out using excitation at 514.5 nm. For signal detection, a CCD detector was employed recording the complete spectral range of interest between 700 and 3,100 cm^{-1}. Adulterations with sunflower oil, rapeseed oil and soybean oil could be easily detected, because the characteristic carotenoid bands to be seen in the spectrum of olive oil at 1,008, 150 and 1,525 cm^{-1} decrease significantly when these edible oils are added. The amount of sunflower oil in adulterated extra-virgin olive oil could be determined by applying the PLS regression algorithm. In this context a lower detection limit of approx. 500 mg/kg adulterated oil was determined.

Usually, the amount of *trans*-isomers in partially hydrogenated vegetable oils is detected by GC or mid-infrared spectroscopy but Raman spectroscopy was also evaluated in comparison to mid IR spectroscopy to predict the content of *cis*- and *trans*-isomers in processed canola and soybean oils [22, 23].

The studies of [24] show that Raman microscopy is useful to discriminate even between different tocopherols and their oxidation products such as α-tocopheryl quinine. Parker and Bisby [1] presented time-resolved resonance Raman spectra of the α-tocopheroxyl radical which represents an intermediate in the antioxidative reaction of α-tocopherol in cellular membranes. The spectra reflect the extent of delocalisation of the radical site due to interaction with the ion pair of the *para*-oxygen atom.

The epicuticular wax of mature mango fruits (*Mangifera indica*) were successfully characterised by means of Raman spectroscopy [25]. Raman measurements presented two morphologically different parts in the wax layer, a crystalline outermost layer and a smooth layer next to the cuticula. Furthermore, the Raman studies were useful to clarify problems faced by the mango industry related to discoloration around lenticels.

Table 9.1 Raman vibrational modes resulting from plant amino acids and proteins. [10, 28]

Analyte	Wavenumber (cm^{-1})	Vibrational mode
Cystine	510	S-S stretch
Cysteine	525	S-S stretch
Methionine	630–670	C-S stretch
	700–745	C-S stretch
	2,550–2,580	S-H stretch
Tyrosine	850/830	Fermi resonance between ring fundamental and overtone
Tryptophan	760, 880, 1,360	Indol ring
Phenylalanine	1,006	Ring breathe
Histidine	1,409	N-deuteroimidazole
Aspartic and	1,400–1,430	C=O stretch of carboxyl group
Glutamic acid	1,700–1,750	C=O stretch of carboxyl or ester group
Amide I	1,655–1,685	Amide C=O stretch, N-H Wagging
Amide III	1,235–1,280	N-H in-plane bend, C-N stretch

9.2.2 Proteins

Numerous studies have been performed to obtain Raman data of plant proteins and amino acids. Some examples of Raman modes useful in the interpretation of protein structure are listed in Table 9.1. Especially sulphur containing amino acids such as cystine, cysteine and methionine show intensive Raman bands in the area of 500–750 cm^{-1} (S-S and S-H stretching bands) as well as in the range between 2,550 and 2,580 cm^{-1} (S-H stretching bands). The peptide bond of proteins possess several distinct vibrational modes, of which the amide I and amide III bands have the main use for characterisation of secondary structure [26].

Séné et al. [27] demonstrated that FT Raman spectroscopy is a powerful tool to investigate plant cell walls and their components. They found that the obtained Raman spectra were diagnostic primarily for phenolic materials whereas IR spectra showed higher sensitivity for pectin and protein. The authors stress that Raman spectroscopy in combination with useful chemometrical algorithms may be a powerful alternative for rapid taxonomic classification of different plant species.

Raman spectra of non-feruloylated and feruloylated wheat arabinoxylans show specific bands at 896, 985, 1,278–1,462 and at 1,091–1,123 cm^{-1} assigned to various vibrational modes of polysaccharides. The presence of β-(1→4) glycoside linkages of xylan backbone was seen at 896 cm^{-1}. Signals occurring in the range between 500 and 600 cm^{-1} were identified as vibrations arising from coupled modes of heavy atoms, C–C and C–O stretching [29]. Beside these signals specific marker bands were observed at 1,598 and 1,628 cm^{-1} demonstrating the presence of ferulic acid esters.

9.2.3 Carbohydrates

Already in 1990 the special advantages of FT-Raman spectroscopy for studies of various carbohydrates have been described [30]. The author measured several pure mono- and disaccharides but also polysaccharides such as woody cotton wool and cypressus wood tissues which show the characteristic signals of cellulose, lignin and hemicelluloses with little fluorescence. Based on the hierarchical cluster analysis of Raman spectra obtained from various species of the Apiaceae family, a taxonomical discrimination of different genbank accessions could be achieved [31]. Main spectral influences of the analysed single seeds were assigned to fatty acids, polysaccharides, proteins, and lignin.

9.3 Secondary Substances

Secondary plant metabolites are organic compounds that are not directly involved in the biochemical processes related to growth, development, or reproduction of the organism. Frequently, they play an important role in prevention of cancer, anti-inflammatory and antiviral effects, decreasing cholesterol levels and blood pressure regulation. Specific types of fruits and vegetables are credited with a variety of various health benefits.

9.3.1 Phenolic Substances

The chemical group of phenolic compounds comprises a considerable range of plant substances possessing an aromatic ring bearing one or more hydroxyl substituents. They play an important role in the broad field of resistance against various plant diseases. Furthermore, phenolic substances contribute to antioxidative properties as well as to the individual flavour of plant species. Most common (poly)phenolic substances occurring in various plants are represented by flavones, flavonols, flavonol glycosides, flavanon glycosides, anthocyanins, catechins and chalcones. The chemical structure of these flavonoids can be modified by hydroxylation, methylation, acylation and glycosylation resulting in a huge number of different chemical substances which may occur in the individual plant tissue. Other phenolic compounds are biosynthesized in the phenylpropane pathway such as anethole, eugenol, carvacrol and thymol which represent important components of various essential oils. Several of them are widely used in the fragrance & flavour industry or as additives for food or feed products relating to their antibiotic properties. Generally, phenolic substances show intense Raman bands due to ring deformation vibrations occurring between 740 and 760 cm^{-1} [32–34]. Furthermore, medium to strong signals are registered between 1,180 and 1,260 cm^{-1} which are related to out-of phase C-C-O stretching modes [35].

After cellulose, lignin is the second most abundant organic material on earth. In the past, numerous techniques have been used to characterise the lignin structure in the context of cell plant lignification and biodegradation processes [36]. Contrary to IR, Raman analysis of lignin have been performed quite recently [37–39]. The distribution of lignin and cellulose in the cell walls of black spruce wood (*Picea mariana*) was investigated using a Raman spectrometer equipped with a confocal microscope and 633 nm laser excitation [37]. The Raman spectra obtained from sections of black spruce show cellulose bands at 380, 1,098 and 2,900 cm^{-1}, whereas lignin contributed signals at 1,600 as well as 1,650 cm^{-1} [40]. Based on these spectral features the ratio of lignin to cellulose concentration was calculated across the double cell wall. The authors stress that Raman imaging permits analyzing plant cell walls without the necessity to use stain, dye and contrast agents. Recently, the same group developed a new SERS Raman method to increase sensitivity and selectivity of lignin in various lignin-containing materials [41]. They mixed the wood samples in presence of ethanol with silver particles (100–150 nm and 2–3.5 μm) and observed that Raman spectra of milled-wood-lignin were surface-enhanced when compared to the normal FT-Raman spectra. Furthermore, a number of Raman signals are shifted to lower wavenumbers suggesting a complexation with silver particles and consequently a change of frequency in the SERS spectra. Contrary to that, silver-adsorbed cellulose samples did not produce a SERS spectrum. It is assumed that π-electrons of the aromatic ring system interact with the silver particles. Therefore, the signals observed in the SERS spectrum of spruce wood can be predominantly related to lignin vibrational modes [41].

Various plant species used as bioenergy feedstocks (pine, oak, poplar, kenaf, miscanthus, pampas gras, switchgrass, alfalfa, orchard grass, and red clover) were analysed for their lignin guaiacyl and syringyl monomer content applying FT Raman spectroscopy [42]. The interpreted Raman data were correlated with the ^{13}C-NMR spectral dataset; furthermore PCA as well as PCR models were developed for classifying plant samples and quantifying lignin components.

Numerous organic colouring substances belong to the chemical class of flavonoids. Beside anthraquinones and indigoids they were the main source of textile colour until the mid- to late 19th century. The most important plants containing yellow dyes include weld (*Reseda luteola*), old fustic (*Chlorophora tinctoria*) and dyer's greenweed (*Genista tinctoria*) [43]. Several vibrational studies have been performed with flavonoids [44–47] but only a limited number of Raman spectra of natural dyes are available in literature [48–50] (natural dyes will be also discussed in the next chapter **Chemometric analysis of Raman and IR spectra of natural dyes**). Surprisingly, only few applications of *in situ* and *in vivo* Raman measurements of anthocyanins have been described [51–53]. Recently, model systems containing anthocyanins, metals (aluminium and iron) and pectins were investigated to analyse the interactions in the individual metal chelate complexes [50]. In this context two marker bands were identified (at approx. 1,510 and 1,330/1,350 cm^{-1}) being responsible for the formation of metal chelates with cyaninidin-3-glycoside and delphinidin-3-glycoside as ligands. Pelargonidin-3-glycoside did not form metal chelates, which was clearly proved by missing marker bands (Fig. 9.1).

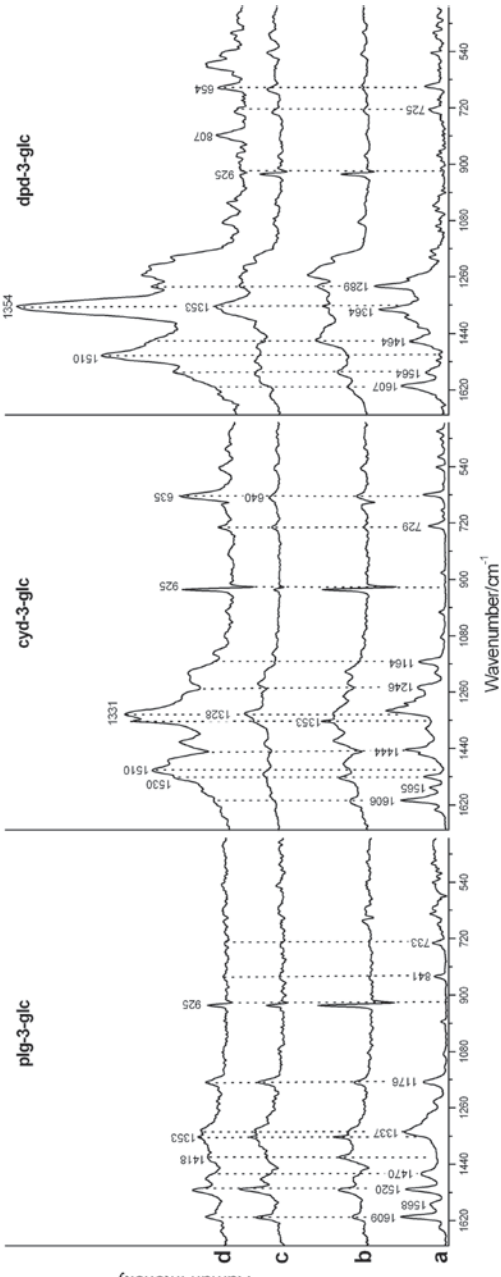


Fig. 9.1 Raman spectra of pelargonidin- (plg-3-glc), cyanidin- (cyd-3-glc) and delphinidin-3-glucosides (dpd-3-glc) as chloride salts (**a**), in acetate buffer solution at pH 5.0 (**b**), with addition of aluminium (**c**) and ferric (**d**) ions in acetate buffer at pH 5.0. Spectra of samples containing buffer were normalised to the CC stretching vibration from acetate at 925 cm^{-1} and calibrated on signals at 2,935 cm^{-1} and 925 cm^{-1}. Then spectrum of acetate buffer has been subtracted. Spectra of chlorides are normalised to the stretching vibration of the phenyl rings at ~1,607 cm^{-1}. [50]

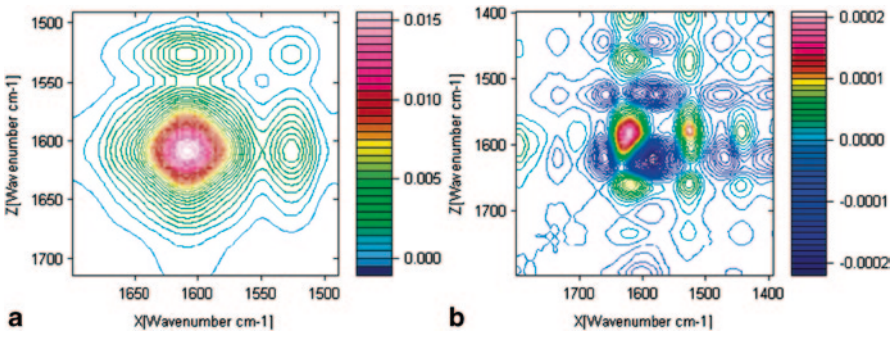


Fig. 9.2 Synchronous (**a**) and asynchronous (**b**) 2D correlation spectra of rooibos. [54]

Furthermore, the absence of anthocyanin-metal chelates in presence of citric acid could be confirmed.

FT-Raman *in situ* measurements of aspalathin, a dihydrochalcone occurring in leaves and stems of unfermented rooibos (*Aspalathus linearis*) has been also performed [54]. Rooibos is indigenous to South Africa and increasingly of interest for the industry due to its nutraceutical properties [55]. The authors demonstrated for the first time that FT-Raman spectroscopy is useful for the identification of aspalathin in dried, green rooibos and the development of a Raman based quantification of aspalathin, nothofagin and total dihydrochalcone levels. Additionally, the advantage of two-dimensional (2D) correlation algorithms for reliable band assignment and an improved Raman quality analysis is presented. In this context a set of rooibos samples with known, increasing concentration of aspalathin was selected, which was used as an external perturbation for creating synchronous and asynchronous spectra. Their counter maps in the wavelength range 1,800–1,400 cm^{-1} are presented in Fig. 9.2. Autopeaks, observed in the synchronous spectrum, represent bands of which the intensities vary most significantly with increasing concentration. The asynchronous 2D correlation maps show a clear symmetric cluster pattern, comprising two positive and two negative cross peaks. This proves that an increase in aspalathin content in the rooibos samples results in an expansion of the band near 1,610 cm^{-1}. It must therefore be associated with higher intensities of those bands assigned to aspalathin, occurring below and above 1,610 cm^{-1} (Fig. 9.2).

Furthermore, by using the FT-Raman mapping technique the spatial distribution of aspalathin in intact rooibos leaves is demonstrated [54].

A few SERS studies have been performed so far with flavonoids [56] and anthraquinone [57, 58]. In most cases the SERS spectra of the individual substances show a very different spectral profile when compared to the normal Raman spectrum due to the chemical interactions with silver colloids. Additional structural information were obtained from the SERS spectra applying different excitation wavelengths (514.5, 632.8, 785 and 1,064 nm), as well as measurements at different pH values.

Very detailed SERS studies have been performed with flavone and three of its hydroxyl derivatives (3-hydroxyflavone, 5-hydroxyflavone and 3,5,7,3′,4′-pentahy-

droxyflavone) [56]. Whereas the SERS spectrum of flavones was weak both on the silver colloid and on the applied silver electrode, the correspondent spectra of the hydroxyl derivatives showed intense bands, indicating that the OH-groups assist the silver colloids to attach on the surface of the analyt molecules. SERS Raman spectroscopy on silver colloids has been also successfully applied to identify yellow dyes occurring in fruits of sumac (*Rhus* sp.) and pomegranate (*Punica granatum*) [59]. The SERS spectrum of sumac extract shows a very good correspondence with the hydroxyflavylium structure of anthocyanins. According to the literature bands registered at 1,637, 1,585, 1,534, 1,485, 1,430 and 1,357 cm^{-1} were assigned to ring stretching vibrational modes, whereas the band at 1,247 cm^{-1} was attributed to the C-O stretching mode [44, 45, 52]. Contrary to that, the aqueous extract of pomegranate rind shows predominantly the spectral feature of ellagic acid, indicating that there is a strong interaction between this phenolic plant substance and the silver colloids. Consequently, a general shift of stretching vibrations related to the aromatic ring system to lower wavenumbers is observed.

Raman spectra of hydroxycinnamic and hydroxybenzoic esters (caffeates, ferulates and gallates), possessing antioxidant and anticarcinogenic properties, with special focus on the analysis of the ring substituents and the nature of the ester alkyl chain was performed [60]. The individual vibrational pattern was assigned in the light of DFT calculations resulting in a good agreement between calculated and experimental data sets.

Two SERS procedures were developed for the evaluation of alizarin in madder root (*Rubia tinctorium*), which usually shows high fluorescence during normal Raman measurements [58]. The first procedure involves thermal evaporation of silver nanoparticles on filter paper stained with the individual natural dye to induce the SERS effect. For the other procedure a special silver-alumina-based substrate (Ag-Al_2O_3) was prepared by spin-coating an alumina nanoparticle layer onto a glass slide to provide the nanostructure of the substrate, followed by thermally evaporating a layer of silver on top of the alumina layer. The dye samples to be analysed were directly delivered onto this substrate. The authors found that the Ag-Al_2O_3 substrate proved to present numerous advantages over other SERS materials and guaranteed a certain surface roughness necessary to induce a strong SERS effect.

A fast and reliable quantification method for the determination of naturally occurring indigo dye in dyer's knotweed (*Polygonum tinctorium*) based on FT-Raman spectroscopy and reference HPLC data was developed [61, 62]. The samples used for the spectroscopic calibration model were prepared by "diluting" pure indigo with cellulose in the range from 0.1 to 2.0 mass %. For this, the indigo was carefully mixed with pure cellulose and 0.2 g of this mixture was used to prepare sample pellets of distinct thickness and diameter by a commercial manual hydraulic press. Alternatively, a conventional sample preparation was performed by pressing a certain amount of the indigo-cellulose mixture in small aluminium pans. Both sample preparation techniques show only small differences in prediction quality. Generally, an excellent correlation within the regarded concentration range could be achieved resulting in correlation coefficient values of 0.997 (metal pan) and 0.988 (pressed pellets).

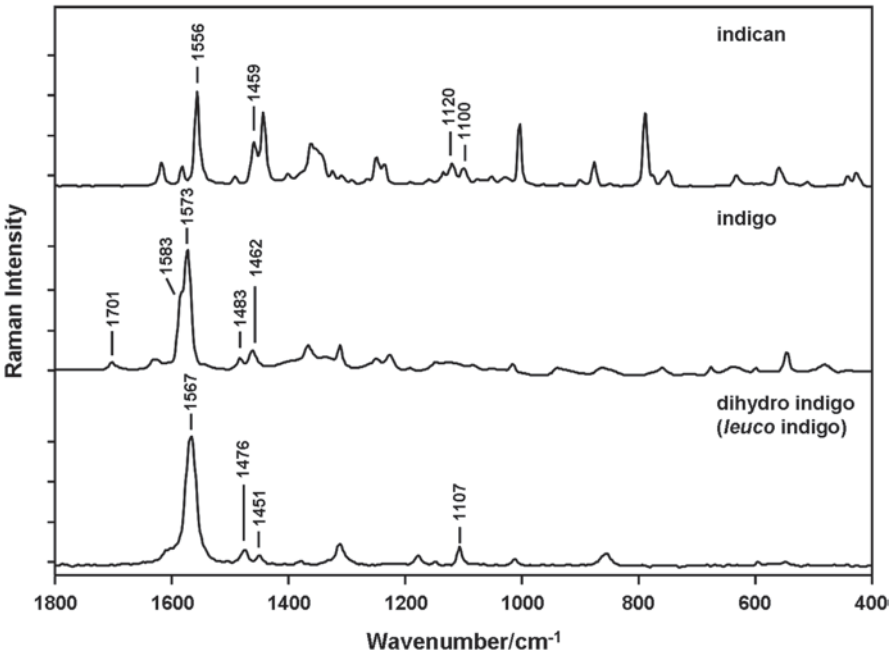


Fig. 9.3 Comparison of the Raman spectra of indican (indoxyl-*β*-D-glucoside), indigo and *in situ* synthesized dihydro indigo. [61]

Besides the qualitative aspects of signal assignment and comparison to appropriate ATR-FT infrared measurements, the Raman spectrum of dihydro indigo, an important intermediate in the indigo dying process, is presented for the first time and discussed with regard to its spectroscopic behaviour. During textile dying, the insoluble blue indigo is reduced to the yellow coloured leuco indigo, which immediately oxidizes back to indigo in the presence of air oxygen [61]. This oxygen-sensitive intermediate of indigo reduction allows the bonding to the dyed strands and hence causes the colour fastness of the textile. The Raman spectra of *in situ* produced leuco indigo, and the observed spectral shifts are shown in Fig. 9.3.

As can be seen, the carbonyl stretch vibration $\nu_{C=O}$ present in indigo around 1,701 cm^{-1} disappears and a new signal around 1,107 cm^{-1} referring to ν C-OH is observed in the spectrum of the *leuco* form. The authors explain the spectral changes by the reduction of the carbonyl oxygen of indigo to the corresponding alcohol in the dihydro indigo. Similar spectral features are observable in the spectrum of indican, where the oxygen forms the acetal bond to the glucose moiety. Together with the stretch vibrations of the sugar C-O bonds, these groups explain the bands around 1,120 and 1,100 cm^{-1} [61].

Normal dispersive Raman and SERS spectra of cape jasmine (*Gardenia augusta*) have been measured by Cañamares et al. [57]. Most intensive bands were found at 1,537, 1,209 and 1,165 cm^{-1} which are mainly related to the natural yellow

components *trans*-crocetin and *trans*-crocin occurring in the plant tissue. Best normal Raman spectra were obtained applying excitation at 785 nm providing highest signal intensity of the dye bands.

9.3.2 Alkaloids

Numerous alkaloids, showing a large range of different chemical structures of which most possess a significant biological activity, have been analysed by Raman spectroscopy [63]. Especially caffeine, theophylline and theobromine show characteristic strong features between 1,310 and 1,340 cm^{-1} and between 1,690 and 1,710 cm^{-1}, which are assigned as C-N and C=N stretching, respectively [64]. The three purine alkaloids have been detected in cacao seeds and their extracts applying FT-Raman spectroscopy [65]. In this context it was also found that theobromine seems to occur in its free form in cacao seeds because no spectral shift was registered compared to the pure standard.

Main alkaloids occurring in poppy (*Papaver somniferum*) and in related pharmaceutical products were successfully characterised by Raman spectroscopy [66]. In the fingerprint range between 700 and 1,500 cm^{-1} various sharp bands can be identified, which are mainly due deformation and stretching vibrations of the alkaloid ring system. Generally, morphine, codeine, and thebaine show no significant differences in their Raman spectral profile, however, some characteristic key bands can be found in the ranges of 1,600–1,650 cm^{-1} and 630–650 cm^{-1} which may be used to discriminate the four mentioned isoquinoline alkaloids.

Raman spectra obtained from poppy milk present characteristic well resolved morphine bands (peaks at 631, 1,620 and 1,642 cm^{-1}). Furthermore, the alkaloids occurring in poppy capsules (morphine, codeine, papaverine, and noscapine) could be reliably predicted. The total alkaloid content of poppy samples was determined with comparatively high prediction quality as well ($R^2=0.90$) [66].

Quinine is the main alkaloid present in the bark of different *Cinchona* species (Rubiaceae), especially *Cinchona pubescens*. The bark of trees of the *Cinchona* genus is the source of a variety of alkaloids, the most familiar of which are the following pairs of stereoisomers: cinchonine and cinchonidine, quinine and quinidine, as well as dihydroquinidine and dihydroquinine. Quinine and quinidine have a bitter taste and due to this property they are considerably used as food additives. The high potential of resonance Raman spectroscopy was demonstrated by the detection of small amounts of these alkaloids in plant materials, without any further sample preparation. Non-resonant Raman microscopy with excitation wavelengths in the visible (532 and 633 nm) and NIR (1,064 nm) regions was unsuccessful due to a strong fluorescence background and due to noisy signals resulting from low laser power, respectively. The use of higher laser power resulted in destruction of the plant material. Non-resonant Raman spectroscopy suffers from low scattering cross section, but sometimes for the active agents which are present at low concentration in the biological environment the signals can be strongly enhanced if the excitation

laser lies within an electronic absorption of the sample. So when UV resonance Raman microspectroscopy (λ_{exc} = 244 nm) was applied for analysis of quinine in *Cinchona* bark the methodology allowed Raman spectra of the active agents to be obtained selectively, with enhanced signals by a factor of up to 10^6 compared to the non-resonant Raman signal of that plant. It was found that the UV resonance Raman spectrum obtained from *Cinchona* bark corresponds well with that of pure isolated quinine. Moreover, in situ UV-RR microspectroscopy is capable of differentiating between quinine and its diastereomer quinidine, structurally very similar active agents. This discrimination is possible based on a marker band at 831 cm^{-1} in the Raman spectrum of quinine, which is shifted to 843 cm^{-1} in the case of quinidine. This vibration involves a banding motion within the side chain around the chiral centre of quinine. Vibrations belonging to the quinoline ring, important for its antimalarial activity in forming p–p interactions to hemozoin, and the vinyl group are resonantly enhanced in the UV-Raman spectra. The interpretation of the experimental spectra was based on DFT calculations combined with FT-Raman spectroscopy of the pure isolated standards. The solvent effect was correlated with the shift of the band at 1,362 cm^{-1} to 1,371 cm^{-1} for anhydrous quinine. This vibration is sensitive to the presence of an aqueous environment and is assigned to a C=C stretching mode [67].

Some other antimalarial alkaloids (dioncophylline A, dioncophylline C, and dioncopeltine A) belonging to the group of naphthylisoquinoline alkaloids were isolated from the tropical liana *Triphyophyllum pelatatum* [68]. In order to locate those parts of the plant containing the highest concentration of these bioactive substances, Raman microspectroscopy was found to be a very efficient tool capable of differentiating between various structurally similar naphthylisoquinoline alkaloids [69]. In this context, the signals registered at 1,356 and 1,613 cm^{-1} assigned to C=C stretching and C–H bending vibrations were especially useful for the reliable distinction of the different alkaloid structures [70]. Raman spectra obtained from the individual pure alkaloid standards, as well as *in vivo* measurements on the plant tissue, were presented and discussed in detail. Most of the identified signals could be successfully assigned to various vibrational modes of the alkaloid molecular structures.

9.3.3 Essential Oils

Raman spectra obtained from essential oils also show characteristic key bands which can be used as marker bands to discriminate different plant species, cultivars and chemotypes. Characteristic group frequencies of numerous pure essential oil components have been listed to support the identification of unknown samples [34, 35, 71–74].

In some cases chemotaxonomic discrimination of plant species is possible measuring the individual fresh plant tissue [3, 75]. As can be seen in Fig. 9.4 micro-Raman measurements performed *in situ* on the essential oil cells of *Eucalyptus cinerea* leaves provide an impression of the individual essential oil distribution [76].

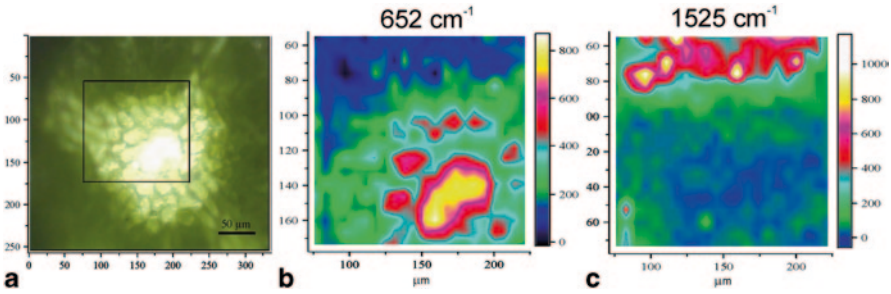


Fig. 9.4 Microscopic image of a leaf section of *Eucalyptus cinerea* presenting the essential oil cavity (**a**) and Raman maps obtained from the defined area colored according to the intensity of the band at 652 cm^{-1} (**b**) and 1,525 cm^{-1} (**c**) showing the distribution of essential oil (b) and carotenes (c). [76]

Eucalyptol (1,8 cineol), representing the main component of the essential oil, presents a strong ring vibration at 652 cm^{-1} which can be therefore used as spectroscopic marker. Spectra taken from the leaf matrix show significant bands at 1,525 and 1,157 cm^{-1} which were assigned as in-phase C=C (ν_1) and C-C stretching (ν_2) vibrations related to the polyene chain of carotenoid vibration modes [76]. Raman image obtained from the integration of the signal at 1,525 cm^{-1} (Fig. 4 C) shows the distribution of carotenoids in the measured area. As expected maximum amounts of carotenoids can be seen outside the essential oil cavities.

Similarly, FT-Raman microspectroscopy of fennel fruits have been performed, revealing the distribution of anethole, the main substance of fennel oil [77, 78]. *In situ* studies of the essential oil cells show two characteristic marker bands at 1,657 and 1,609 cm^{-1} (ring stretching modes of anethole) which were used for the Raman mapping measurements.

New SERS Raman methods have been developed to perform *in situ* investigation of the essential oils of *Mentha piperita* [75] and *Thymus vulgaris* [3]. The SERS spectrum of *T. vulgaris* presents a strong Raman signal at 738 cm^{-1} due to ring stretching of thymol. Contrary to the corresponding normal Raman spectrum signals of carvacrol and p-cymene were not identified probably due to their low sample concentration or less affinity towards silver ions. Silver-coated glass fibre tips have been applied for the investigation of the essential oils of *Mentha piperita* and *M. citrata* directly in the glandular trichomes on the plants. The "tip" spectra show significantly improved quality compared to the corresponding SERS spectra also providing additional spectral information [79].

Based on the fingerprint region of the individual Raman spectra the main components of pepper oil can be clearly identified [80]. The Raman spectrum of the so-called "caryophyllene type" shows comparatively strong stretching vibration modes at 1,632 and 1,671 cm^{-1} and a broad band at 1,446 cm^{-1} which is assigned to a CH_2-deformation vibration. Also the Raman spectrum of the "sabinene/caryophyllene type" presents these marker bands of caryophyllene but additionally the characteristic bands of sabinene ($\nu_{C=C}$ at 1,653 cm^{-1} and $\delta_{CH2\ at}$ 1,415 cm^{-1}) can be

seen. Furthermore, several sabinene bands of minor intensity occur in the range between 600 and 1,000 cm^{-1} (δ_{C-C} at 950 cm^{-1} and ring deformation vibration at 652 cm^{-1}). The Raman spectrum of the "δ-3-carene/caryophyllene/limonene type", shows signals of all three main components, δ-3-carene (key bands at 1,686, 819, 713 and 678 cm^{-1}), caryophyllene (key band at 1,632 and 1,671 cm^{-1}) and limonene (key band at 760 cm^{-1}).

Raman spectroscopy was also successfully applied for chemotype discrimination of marjoram and oregano essential oils as well as for identification of the essential oil constituents [81]. The essential oils isolated from various basil species and cultivars have been also analysed by means of Raman spectroscopy and 11 chemotypes could be successfully classified [82].

Raman spectra obtained from various citrus oils (orange, grapefruit, mandarin, lemon, and lime) [72] show more or less a similar spectral profile. Some of the most relevant monoterpene components occurring in the oil, such as limonene and γ-terpinene, show characteristic key bands, which are useful for a quantitative distinction of the different oil types. Applying principal component analysis (PCA) to the spectral data, the individual citrus oils can be clearly separated in a 3-dimensional factor space. Cross-validation statistics showed that a partial least square (PLS) algorithm was able to produce high quality predictions of the relative oil composition as well as general quality parameters. The dominating bands are due to limonene vibrations and are recorded at 760, 1,435, 1,644/1,678 cm^{-1}. Based on the first three PCA factors, 86 % of the total variation could be explained. Applying the PLS algorithm all relevant oil components as well as sum parameters (total aldehyde content, optical rotation) of the various citrus oils were precisely analyzed.

Calibration curves for the prediction of α-pinene and β-myrcene occurring in mastic gum oil have been developed [74]. The authors report, that the correlation coefficient was calculated to be 0.992 and 0.997, respectively.

Recently, a novel approach was reported to analyse in situ (-)-bornyl acetate in the essential oil of Sibirian fir (*Abies sibirica*) by means of Raman optical activity (ROA) [83]. Characteristic Raman signals were observed at 1,688, 1,659 and 1,639 cm^{-1} due to C=C stretching vibrations originating from the four main components (bornyl acetate, camphene, α-pinene, 3-carene) of the essential oil. The experimental results correlated very well with simulated ROA and Raman spectra of bornyl acetate obtained from DFT calculations. The authors mention that this method has the potential to become a standard procedure for commercial essential oil analysis.

9.3.4 Carotenoids

Carotenoids occur in chromoplasts of plants and algae and are responsible for their characteristic yellow, orange or red-color. Generally, plant pigments such as carotenoids, anthocanins and flavonoids fulfill the task to attract insects and birds for the purpose of pollination and thus to guarantee the preservation of the individual

species [84]. The main role of carotenoids in plants is related to photosynthesis where they interact between light absorption, excitation energy transfer, dissipation of excess energy and electron transfer processes. Carotenoids are located in antenna-proteins forming light-harvesting complexes, which surround a photosynthetic center to focus photon energies absorbed by the pigment, and to transport the excitation energy to the reaction center where it is converted into chemical energy driving the photosynthetic reaction [85].

Raman spectra of β-carotene show three characteristic bands at 1,524, 1,157 and 1,005 cm^{-1}, respectively, which are assigned as double bond C=C stretching mode (1,524 cm^{-1}), C-C in-plane single bond stretching mode (1,157 cm^{-1}), and C-H bending mode (1,008 cm^{-1}), respectively.

Already in early 1970s the resonance Raman spectra for astaxanthin were measured and a correlation between the absorption spectra and the resonance Raman excitation profiles of this carotinoid substance was found [86].

Recently, it has been observed that the position of C=C stretching vibrations in the Raman spectrum is mainly influenced both by the carotenoid polyene chain length and the structure of terminal substituents [87]. Based on this finding it is principally possible analyzing simultaneously the distribution of different carotenoids in plant tissue [88]. Even unstable epoxycarotenoids such as auroxanthin and violaxanthin were detected with high sensitivity and their distribution in the plant tissue could be measured *in situ* applying micro-Raman mapping [53, 88]. Furthermore, FT-Raman microspectroscopy has been used to demonstrate the microstructure and chemical composition of fennel fruits, chamomile inflorescence and curcuma roots [77]. FT-Raman imaging or mapping of carotenoid distribution also seems to be a suitable approach to show possible changes in the plant metabolism induced by abiotic and biotic stresses as well as sunscald physiological disorder [89].

Changes of carotinoid distribution was also found in leaves of diseased sugarbeet (*Beta vulgaris* L.) and in leaves of parsley (*Petroselinum crispum* Mill. Nym.) which were infected by the fungus *Septoria petroselini* [89].

It is well known that during ripening the composition of the carotenoid fraction sometimes changes dramatically. Also in this context Raman spectroscopy was successfully applied e.g. to characterize the content of lutein, β-carotene and capsanthin in the individual ripening stages of bell pepper (*Capsicum annum*) [89]. Quite recently, confocal Raman has been used to identify in parallel lycopene, β-carotene, and lutein in tomato fruits [90]. The authors mention that this technique provides some advantages for the food industry in order to get a rapid information of fruit quality during processing.

Comprehensive Raman studies of carrots (*Daucus carota*) show that various quality parameters such as carotenoids, polyacetylenes, pectin and starch can be determined simultaneously in wild and cultivated roots [90, 92]. Moreover, the distribution of α- and β-carotene as well as lutein and lycopene was analysed in root sections of various carrot cultivars. This finding was possible due to the ability to separate different Raman bands to β-carotene (1,520 cm^{-1}), lycopene (1,510 cm^{-1}) and α-carotene/lutein (1,527 cm^{-1}). Clearly, the level of β-carotene was heterogeneous across root sections of orange, yellow, red and purple roots. In the second-

ary phloem β-carotene increased gradually from periderm towards the core, but declined fast in cells close to the vascular cambium. On the other hand, α-carotene and lutein were deposited in younger cells with a higher rate than β-carotene, while lycopene in red carrots accumulated throughout the whole secondary phloem at the same level [93].

Reliable FT-Raman methods were developed for the rapid determination of lycopene and other carotenoids occurring in tomatoes and various related products such as tomato puree or tomato ketchup [94]. Compared to the spectrum of pure isolated lycopene, the most intense C=C stretching vibration of this substance was found to be slightly different in the Raman spectra obtained from the plant samples. The authors assume that this shift is mainly caused by the interaction between the naturally occurring carotenoids and the plant matrix changing their chemical form or physical structure.

Using an FT-Raman spectrometer equipped with a microscope (40-fold amplification) carotenoid crystals sequestered in a carrot cell have been also measured at the subcellular level [95]. The characteristic bands occurring at 1,518 cm^{-1} and 1,156 cm^{-1} proved that the crystals were predominantly formed from β-carotene.

Applying nonlinear coherent anti-Stokes Raman scattering (CARS) microscopy measurements of carotenoids in various thermally treated plant tissues such as sweet potato, carrot, and mango were performed [96]. Based on the obtained CARS signals, a quantification of size, shape, density, and location of different carotenoid bodies could be achieved. In this context heterogeneous rod-shaped bodies with high carotenoid densities were found for potatoes and carrots, whereas the carotenoid containing lipid droplets inside the mango fruits presented a more homogeneous low-density and rounded shape. After thermal treatment of the potatoes, β-carotene density and morphology remained intact despite of significant changes of the surrounding starch granules [96].

A range of naturally occurring carotenoids in numerous plant species have been analysed using FT-Raman spectroscopy [97]. Here the authors stress that there may be a significant wavenumber shift caused by carotenoid interactions with the surrounding plant tissue. The special potential of Raman spectroscopy was also applied to characterize the different ripening stages of olives [98]. The authors showed that the increase of triglycerides, carotenoids and phenolic compounds during the first development stages and their decrease during the ripening process can be successfully monitored by means of the Raman bands at 1,525 and 1,605 cm^{-1}, respectively.

Raman spectroscopy has been also applied to discriminate pollen grains of different tree species [99–101]. In this context single pollen Raman spectra are presented showing the individual key bands of carotenoids, proteins, nucleic acids, carbohydrates, and lipids [99]. The classification of the different pollen species was based on PCA as well as hierarchical cluster analyses (HCA) and allowed to analyze various chemical classes of molecules simultaneously [100]. A similar study aiming to measure the individual carotenoid profiles of different pollen grains of trees was performed using the strong pre-resonant Raman spectra obtained before and after photodepletion of the carotenoid molecules [101]. Without any further sample pre-

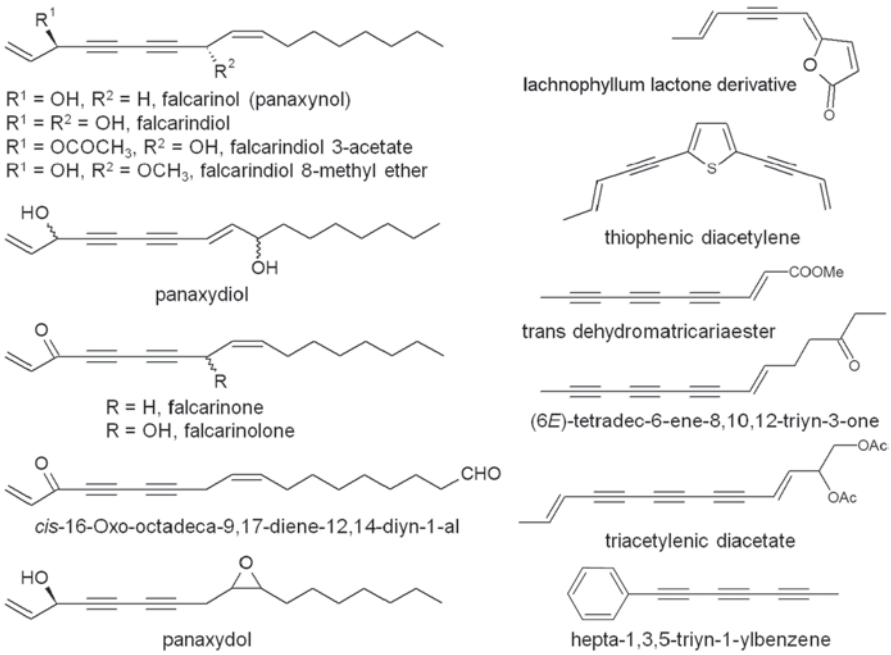


Fig. 9.5 Structures of falcarinol-type diacetylenes and other exemplary mono- and polyacetylenes. [103]

treatment, carotenoids occurring in single pollen grains were examined achieving both, high detection sensitivity and structural selectivity.

The special advantages of Raman spectroscopy over classical chemical analysis methods especially for industrial purposes were pointed out by [102]. Here, a rapid and inexpensive quantitative Raman analysis is presented for the determination of carotenoids in inflorescences of *Tagetes erecta*. Other authors mention that Raman spectroscopy is a very useful technique for rapid and reliable characterization of newly isolated plant carotenoids.

9.3.5 Polyacetylenes

Most common polyacetylenes occurring in the plant kingdom are aliphatic acetylenes. Especially, falcarinol-type acetylenes with two conjugated triple bonds in their structures are widely distributed in the *Apiaceae* and *Araliaceae* plant families [104]. Today, more than thousand different polyacetylenes have been detected in nature (Fig. 9.5) [105]. Falcarinol-type polyacetylenes, biosynthesised from oleic acid, represent bioactive substances protecting against pest infestation caused by insects, bacteria and fungi. Polyacetylenes are also effective as phytoalexins. Furthermore, a stimulation of polyacetylenes by fungal attack has been observed in several plant families such as *Asteraceae*, *Leguminosae*, and *Apiaceae* [106]. Individual

polyacetylenes occurring in the plant tissue can be detected by means of FT-Raman spectroscopy even in comparatively low amounts of 0.01–0.1 g/100 g fresh matter. Raman spectra of diacetylenes show a very strong and polarized band due to -C≡C-C≡C- symmetric stretching vibration in the region of 2,200–2,300 cm^{-1} [71, 107, 108]. It has been observed that both, the number of conjugated -C≡C- bonds in polyacetylenes and substituents contribute to the number of Raman signals, their frequencies and intensities [71]. Therefore, the individual shape of the region around 2,200 cm^{-1} provides first informations with regard to the individual polyacetylene type. Thus, the diacetylene substances falcarinol, falcarindiol, and panaxydol present similar characteristic bands which are shifted by only a few cm^{-1}. The band maximum for falcarinol appears at 2,258 cm^{-1} and for falcarindiol at 2,252 cm^{-1} (caused by the additional -OH group in α-position of the diacetylenic system), whereas the Raman spectrum of panaxydol shows the corresponding band at 2,260 cm^{-1} [71]. Contrary to that, Raman spectra of unsubstituted triacetylenes show two strong bands in the region between 2,190 and 2,120 cm^{-1}, due to in-phase and out-of phase stretching vibrations of all triple bonds [71]. It has been found that tetraacetylenes show two symmetric and antisymmetric -C≡C- modes, whereas diacetylenes substituted by a group with a similar mass as that of the -C≡C- moiety, several additional couplings occur and the spectral shape in the ν(C≡C) region becomes more complicated [109].

Generally, polyacetylenes can be reliably quantified even if they occur in lower concentration in plant tissue. This observation is mainly due to the high scatter induced by -C≡C- bonds in a frequency region not interfering with other vibrational modes. Furthermore, it has to be mentioned here that the presence of polyacetylenes in plant tissues can be clearly illustrated by *in situ* Raman mapping.

It has been proved that apart from easy detection of polyacetylenic compounds, FT-Raman spectroscopy is also very useful to distinguish several polyacetylenes [71]. Two similar polyacetylenes, i.e. falcarinol and falcarindiol, usually co-existing in several plants, present a significant wavenumber shift of their characteristic polyacetylene band. Also other polyacetylenes can be distinguished in a similar way. The spectrum taken from the roots of the high-carotene orange carrot (*Daucus carota*) shows a few overlapping bands in the region of 2,260 and 2,250 cm^{-1}. It has been found that the main part of polyacetylenes is located in the outer section of the root, i.e. in the secondary phloem tissue close to the vascular cambium as well as in pericyclic parenchyma tissue close to the periderm [110]. It has been found that the accumulation of polyacetylenes correlates very well with a high concentration of carotenes, but both polyacetylenes and carotenoids occurred at different areas in the investigated plant material.

Wild carrot species (*D. carota* ssp. *maritimus, D. carota* ssp. *gummifer, D. carota* ssp. *commutatus*) have been also investigated regarding the accumulation of polyacetylenes. Raman spectra prove that mainly falcarindiol is present in the individual wild species. Recently, differences in the Raman spectra of two wild carrot species have been described more detailed (*D. carota* ssp. *gummifer, D. carota* ssp. *maximus*) [111]. Generally, the Raman band at 2,258 cm^{-1} characteristic for falcarinol was missing in the spectra obtained from wild carrot roots implying that falcarinol

was either absent or only present in low amounts. Chromatographic measurements performed by Metzger [112] confirmed this assumption. Furthermore, Raman mapping clearly presented the different distribution of polyacetylenes in wild and orange carrots. It could be shown that the whole phloem tissue was rich in polyacetylenes, but the maxima could have been also observed near the pericyclic parenchyma. An analogous distribution of polyacetylenes has been found in roots of other carrot wild species. In recent studies, Roman et al. [111] show that the distribution of polyacetylenes within the phloem tissue is not uniform and areas with high amounts of these compounds may also occur occasionally. In addition, the authors could not find higher concentration of polyacetylenes close to the periderm. Significant amounts of falcarindiol detected in wild carrot species are related to their resistance to some plant diseases such as root knot nematodes [113].

As already mentioned above polyacetylenes mainly occur in plant species from the Apiaceae family. Recently, parsnip (*Pastinaca sativa*), celeriac (*Apium graveolens* var. *rapaceum*), and parsley (*Petroselinum crispum*) have been investigated aiming to show the distribution of the individual polyacetylene substances [114]. The Raman spectra obtained from various areas of the parsnip roots showed distinct signals in the region of 2,180–2,270 cm^{-1} and it has been found that the spectral profiles as well as band positions depend significantly on the localization within the root.

Falcarinol, falcarindiol, 8-O-methylfalcarindiol, panaxydiol, falcarinone and falcarinolone have been detected in celeriac roots applying HPLC, but so far no quantification of these polyacetylenes has been performed. Raman spectra, measured at different areas of the root have presented more or less the same pattern and a band maximum at about 2,252 cm^{-1} [114]. Therefore, the authors assume that falcarindiol dominates in celeriac root whereas the other polyacetylenes occur only in lower amounts. Raman mapping has been also performed to study the polyacetylene distribution within the whole transversely cut root. As to be seen in Fig. 9.6, polyacetylenes are accumulated only in a few small areas occasionally localized close to the peel [114].

In a similar way, Raman spectra taken from parsley roots mainly show characteristic bands related to falcarinol and falcarindiol, but also several smaller signals were found which are due to other polyacetylenes or the response of the plant matrix [114]. A Raman map and the related image of a parsley root is presented in Fig. 9.6 proving a more or less homogeneous distribution of polyacetylenes in the phloem tissue close to the secondary cambium. Polyacetylenes (falcarinol and panaxydol) were also successfully detected in ginseng roots applying Raman spectroscopy. The Raman spectra of the isolated polyacetylenes present signals at about 2,260 cm^{-1}, whereas a significant shift to 2,237 cm^{-1} can be observed when the same compound is measured in the fresh ginseng root. Raman spectra obtained from freeze-dried ginseng roots show a band with a maximum at 2,258 cm^{-1}, while the corresponding polyacetylene signal measured after hydration of the freeze-dried material again presents a band at 2,237 cm^{-1} [93]. The authors explain this phenomenon with an assumed interaction between the free π-electrons located at the polyacetylene bonds and metal ions in presence of water resulting in the formation of a stable metal

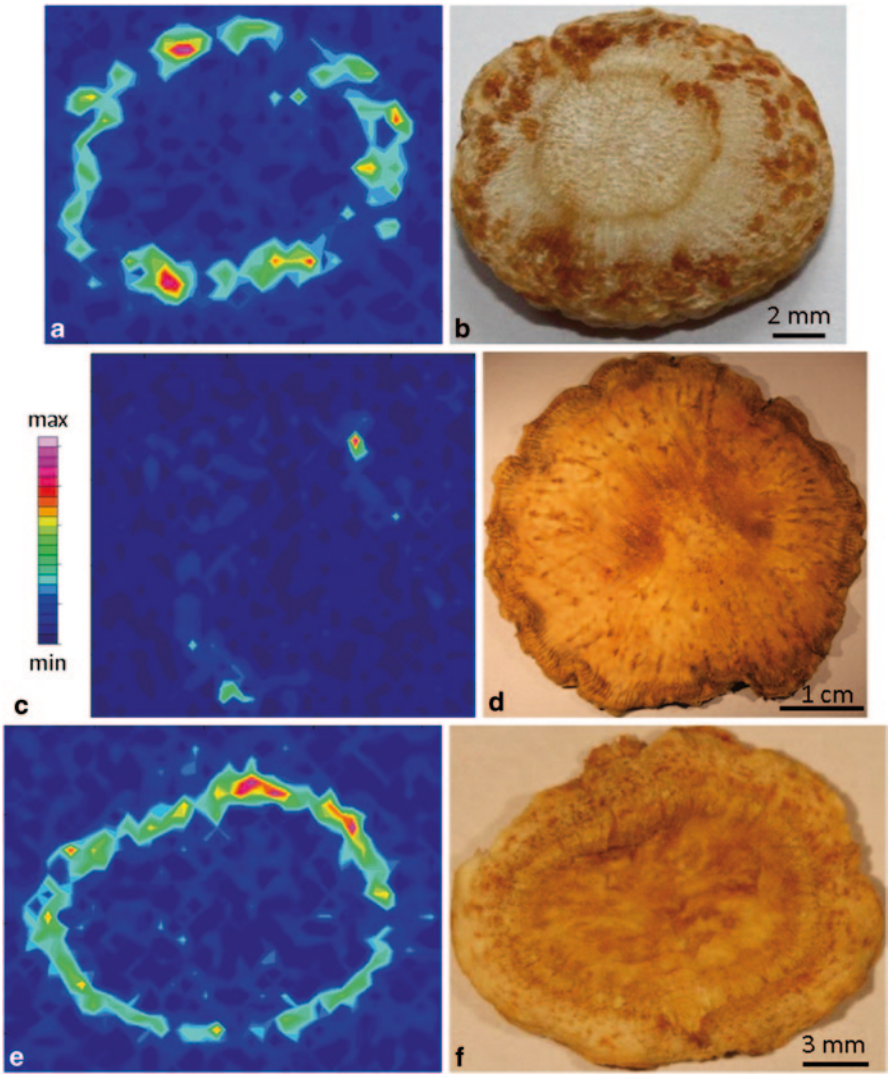


Fig. 9.6 Raman maps and related images obtained from transversely root sections of parsnip (**a**, **b**), celeriac (**c**, **d**), and parley (**e**, **f**). Maps are coloured according to the individual band intensities in the wavelength ranges 2,160–2,280 cm^{-1} (parsnip) and 2,220–2,280 cm^{-1} (celeriac and parsley). [114]

complex. Raman maps obtained from the ginseng roots show a similar distribution of polyacetylenes as formerly registered for carrot roots.

Falcarindiol has been also identified in the roots of Chinese angelica (*Angelica dahurica*) by means of Raman spectroscopy. Here, the characteristic polyacetylene signal was observed at 2,252 cm^{-1}. The Raman spectrum obtained from the flowers of apache beggarticks (*Bidens ferulifolia*) that is mainly grown as horticultural

plant in Guatemala and Mexico, shows similar signals at 2,196 cm^{-1}. Other species of the genus *Bidens* are widely used in Chinese medicine, such as *Bidens pilosa* and *Bidens campylotheca.* Several of them contain lipophilic polyacetylenes as well as polyacetylene glucosides that are predominantly made responsible for the described medicinal properties [115–117]. The distribution of polyacetylenes in the inflorescence of chamomile (*Matricaria chamomilla*) has been also mapped by Raman spectroscopy, integrating the wavenumber range between 2,148 and 2,234 cm^{-1} [66, 108]. The authors were able to demonstrate that the polyacetylenes (*cis*- and *trans*-spiroethers) are mainly present in the bottom of the flower heads.

9.3.6 Miscellaneous

A sensitive SERS Raman method was presented allowing to analyse cyanogenic glycosides in plant tissue [118]. The individual spectra of the cyanogenic substances (linamarin, linustatin, neolinustatin, amygdalin, sambunigrin, and dhurrin) were registered using a Raman microscope equipped with a 532 nm laser. The authors observed the signal of the cyanohydrin group at around 2,240 cm^{-1} being influenced by the nature of side groups in the molecule. In a similar way, the content of the cyanogenic constituent amygdalin in bitter almond (*Prunus amygdalus*) was investigated applying a standard addition series in which different amounts of the analyt (0–700 nmol/mg) were added to ground sweet almond [119]. The best PLS regression model was developed on the vibrational band of the nitrile group (RMSECV: 13 nmol/mg; p: 0.937). Furthermore, the distribution of amygdalin in bitter almond was investigated *in situ* applying Raman microscopy presenting highest amounts in the epidermis. Stellacyanin, a copper-containing glyco-protein occurring in cucumber (*Cucumis sativus*) has been characterised by Resonance Raman spectroscopy [120]. In accordance with earlier Raman studies performed at the Japanese lacquer tree *Rhus vernicifera* [121] the most intense signals in the stellacyanin spectrum were detected in the range between 340 and 430 cm^{-1} and assigned as Cu-S (cysteine) stretching modes.

The triterpene substance betulin, which is obtained from the outer bark of birch tree, is well known for its broad range of biological properties such as anticancer, antibacterial, anti-inflammatory, antifungal and antiviral activities. A complete vibrational assignment of all modes of betulin, based on DFT calculations, is presented by Fălămaş et al. [122]. Also, the individual calculated band positions are compared with the measured FT-Raman data proving that the theoretical model was well computed. Furthermore, betulin and lupeol were successfully identified in a natural extract of birch bark by means of Raman spectroscopy.

Different Raman spectroscopic techniques were used to characterise the three biologically active diterpenoic acids abietic acid, gibberellic acid and kaurenoic acid [123]. Because conventional Raman spectroscopy was not suitable to detect these substances exhibiting very low polarizability, UV-resonance and SERS Raman spectroscopy were applied. Optimal signal sensitivity was achieved with

785-nm excitation on a novel nano-structured substrate coated with a 400-nm gold layer. These spectra showed clear differences and were suitable for the identification of the mentioned diterpenoic acids.

FT-Raman spectroscopy was successfully applied for identification and quantification of harpagoside in secondary roots of devil's claw (*Harpagophytum procumbens*) as well as in related phytopharmaceutical products, e.g., ethanolic extracts and tablets [124]. The obtained Raman maps reveal the spatial distribution of this valuable iridoid glycoside substance within the different plant samples. The authors stress that the same technique can be used for quality control purposes covering the whole process from plant cultivation to the final phytopharmaceutical products. Based on the obtained spectral data and reference HPLC values of harpagoside, a reliable multivariate calibration model was developed [124].

References

1. Parker AW, Bisby RH (1993) J Chem Soc Faraday Trans 89:2873
2. Moskovits M (2006) Surface-enhanced Raman spectroscopy: a brief perspective. In: Kneipp K, Moscovits M, Kneipp H (eds) Surface-enhanced Raman scattering. Springer, Berlin, pp 1–17
3. Rösch P, Popp J, Kiefer W (1999) J Mol Struct 480/481:121
4. Zeiri L (2007) J Raman Spectrosc 38:950
5. Gordea-Torresdey JL, Parson JG, Gomez E, Peralta-Videa J, Troiani HE, Santiago P, Jose-Yacaman M (2002) Nano Lett 2:397
6. Gordea-Torresdey JL, Gomez E, Peralta-Videa J, Parson JG, Troiani HE, Jose-Yacaman M (2003) Langmuir 19:1357
7. Goodacre R, Vaidyanathan S, DunnWB, Harrigan GG, Kell DB (2004) Trends Biotechnol 22:245
8. Kell DB (2004) Curr Opin Microbiol 7:296
9. Baeten V, Hourant P, Morals MT, Aparicio R (1998) J Agric Food Chem 46:2638
10. Li-Chan ECY (1994) Trends Food Sci Technol 5:3
11. Muik B, Lendl B, Molina-Diaz A, Ayora-Canada MJ (2003) Anal Chim Acta 487:211
12. Muik B, Lendl B, Molina-Diaz A, Ayora-Canada MJ (2003) Appl Spectrosc 57:233
13. Heise HM, Damm U, Lampen P, Davies AN, McIntyre PS (2005) Appl Spectrosc 59:1286
14. Reitzenstein S, Rösch P, Strehle MA, Berg D, Baranska M, Schulz H, Rudloff E, Popp J (2007) J Raman Spectrosc 38:301
15. Yang H, Irudayaraj J, Paradkar MM (2005) Food Chem 93:25
16. Barthus RC, Poppi RJ (2001) Vib Spectrosc 26:99
17. Marigheto NA, Kemsley EK, Defernez M, Wilson RH (1998) J Amer Oil Chemists' Soc 75:987
18. Weng RH, Weng YM, Chen WL (2006) J Chinese Chem Soc 53:597
19. Yang H, Irudayaraj J (2001) J Am Oil Chemists' Soc 78:889
20. Omar J, Sarmiento A, Olivares M, Alonso I, Etxebarria N (2012) J Raman Spectrosc 43:1151
21. El-Abassy RM, Donfack P, Materny A (2009) J Raman Spectrosc 40:1284
22. Bailey GF, Horvat RJ (1972) J Amer Oil Chemists' Soc 49:494
23. Johnson GL, Machado RM, Freidl KG, Achenbach ML, Clark PJ, Reidy SK (2002) Organ Process Res Dev 6:637

24. Beattie JR, Maguire C, Gilchrist S, Barrett LJ, Cross CE, Possmayer F, Ennis M, Elborn JS, Curry W, McGarvey JJ, Schock BC (2007) J Fed Am Soc Exp Biol 21:766
25. Prinsloo LC, du Plooy W, van der Merwe C (2004) J Raman Spectrosc 35:561
26. Susi H, Byler DM (1988) Appl Spectrosc 42:819
27. Séné CFB, McCann MC, Wilson RH, Grinter R (1994) Plant Physiol 106:1623
28. Peticolas WL (1995) Methods Enzymol 246:389
29. Barron C, Robert P, Guillon F, Saulnier L, Rouau X (2006) Carbohydr Res 341:1186
30. Goral J (1990) Curr Top Biophys 16:33
31. Baranska M, Baranski R, Schulz H, Nothnagel T (2006) Planta 224:1028
32. Baranska M, Schulz H, Krüger H, Quilitzsch R (2005) Anal Bioanal Chem 381:1241
33. Daferera DJ, Tarantulis PA, Polissiou MG (2002) J Agric Food Chem 50:5503
34. Schulz H, Baranska M (2005) Perfum Flavorist 30:28
35. Lin-Vien D, Colthup NB, Fateley WG, Grasselli JG (eds) (1991) The handbook of infrared and Raman characteristic freqencies of organic molecules. Academic Press Inc, San Diago
36. Lin SY, Dence CW (eds) (1992) Methods in lignin chemistry. Springer-Verlag, Berlin
37. Agarwal UP (2006) Planta 224:1141
38. Vester J, Felby C, Nielsen OF, Barsberg S (2004) Appl Spectrosc 58:404
39. Saariaho A-M, Argyropoulos DS, Jaaskelainen A-S, Vuorinen T (2005) Vib Spectrosc 37:111
40. Agarwal UP, Ralph SA (1997) Appl Spectrosc 51:1648
41. Agarwal UP, Reiner RS (2009) J Raman Spectrosc 40:1527
42. Lupoi JS, Smith EA (2012) Appl Spectrosc 66:903
43. Ferreira ESB, Hulme AN, McNab H, Quye A (2004) Chem Soc Rev 33:329
44. Cornard JP, Vrielynck L, Merlin JC, Wallet JC (1995) Spectrochim Acta Part A 51:913
45. Cornard JP, Merlin JC, Boudet C, Vrielynck L (1997) Biospectroscopy 3:183
46. Vrielynck L, Cornard JP, Merlin JC, Lautie MF (1994) Spectrochim Acta Part A 50:2177
47. Erdogdu Y, Unsalan O, Sajan D, Gulluoglu MT (2010) Spectrochim Acta Part A 76:130
48. Bell IM, Clark JHC, Gibbs PJ (1997) Spectrochim Acta Part A 53:2159
49. Burgio L, Clark RJH (2001) Spectrochimica Acta Part A 57:1491
50. Buchweitz M, Gudi G, Carle R, Kammerer DR, Schulz H (2012) J Raman Spectrosc 43:2001
51. Brouillard R (1983) Phytochemistry 22:1311
52. Merlin JC, Statoua A, Brouillard R (1985) Phytochemistry 24:1575
53. Gamsjäger S, Baranska M, Schulz H, Heiselmayer P, Musso M (2011) J Raman Spectrosc 42:1240
54. Baranska M, Schulz H, Joubert E, Manley M (2006) Anal Chem 78:7716
55. Joubert E, Winterton P, Britz TJ, Gelderblom WCA (2005) J Agric Food Chem 53:10260
56. Teslova T, Corredor C, Livingstone R, Spataru T, Birke RL, Lombardi JR, Canamares MV, Leona M (2007) J Raman Spectrosc 38:802
57. Cañamares MV, Leona M, Bouchard M, Grzywacz CM, Wouters J, Trentelman K (2010) J Raman Spectrosc 41:391
58. Chen K, Leona M, Vo-Dinh K, Yan F, Wabuyele MB, Vo-Dinh T (2006) J Raman Spectrosc 37:520
59. Bruni S, Guglielmi V, Pozzi F, Mercuri AM (2011) J Raman Spectrosc 42:473
60. Calheiros R, Machado NFL, Fiuza SM, Gaspar A, Garrido J, Milhazes N, Borges F, Marques MPM (2008) J Raman Spectrosc 39:95
61. Fiedler A, Baranska M, Schulz H (2011) J Raman Spectrosc 42:551
62. Baran A, Fiedler A, Schulz H, Baranska M (2010) Anal Methods 2:1372
63. Baranska M, Schulz H (2009) Determination of alkaloids trough infrared and Raman spectroscopy. In: Cordell GA (ed) The alkaloids, Vol. 67. Elsevier, Netherlands, pp 217
64. Gunasekaran S, Sankari G, Ponnusamy S (2005) Spectrochim Acta A 61:117
65. Edwards HGM, Villar SE, de Oliveira LFC, Le Hyaric M (2005) Anal Chim Acta 538:175
66. Schulz H, Baranska M, Quilitzsch R, Schütze W (2004) Analyst 129:917
67. Frosch T, Schmitt M, Noll T, Bringmann G, Schenzel K, Popp J (2007) Anal Chem 79:986

68. Bringmann G, Rübenacker M, Vogt P, Busse H, Assi LA (1992) Phytochemistry 31:4019
69. Frosch T, Schmitt M, Schenzel K, Faber JH, Bringmann G, Kiefer W, Popp J (2006) Biopolymers 82:295
70. Frosch T, Schmitt M, Popp J (2007) J Phys Chem B 111:4171
71. Schulz H, Baranska M (2007) Vib Spectrosc 43:13
72. Schulz H, Schrader B, Quilitzsch R, Steuer B (2002) Appl Spectrosc 56:117
73. Schrader B (1995) Infrared and Raman spectroscopy methods and applications. VCH, Weinheim
74. Daferera DJ, Pappas C, Tarantulis PA, Polissiou MG (2002) Food Chem 77:511
75. Rösch P, Kiefer W, Popp J (2002) Biopolymers 76:358
76. Baranska M, Schulz H, Reitzenstein S, Uhlemann U, Strehle MA, Krüger H, Quilitzsch R, Foley W, Popp J (2005) Biopolymers 78:237
77. Baranska M, Schulz H, Rösch P, Strehle MA, Popp J (2004) Analyst 129:926
78. Strehle MA, Rösch P, Baranska M, Schulz H, Popp J (2005) Biopolymers 77:44
79. Gessner R, Rösch P, Kiefer W, Popp J (2002) Biopolymers 76:327
80. Schulz H, Baranska M, Quilitzsch R, Schütze W, Lösing G (2005) J Agric Food Chem 53:3358
81. Schulz H, Özkan G, Baranska M, Krüger H, Özcan M (2005) Vib Spectrosc 39:249
82. Schulz H, Schrader B, Quilitzsch R, Pfeffer S, Krüger H (2003) J Agric Food Chem 51:2475
83. Chruszcz-Lipska K, Blanch EW (2011) J Raman Spectrosc 43:286
84. Mlodzinska E (2009) Acta Biol Cracoviensia Ser Bot 51:7
85. Bode S, Quentmeier CC, Liao P-N, Hafi N, BarrosT, Wilk L, Bittner F, Walla PJ (2009) PNAS 106:12311
86. Salares VR, Mendelsohn R, Carey PR, Bernstein HJ (1976) J Phys Chem 80:1137
87. Gierlinger N, Schwanninger M (2007) Spectrosc Int J 21:69
88. Schulz H, Baranska M, Baranski R (2005) Biopolymers 77:212
89. Baranski R, Baranska M, Schulz H (2005) Planta 222:448
90. Pudney PDA, Gambelli L, Gidley M (2011) Appl Spectr 65:127
91. Baranski R, Baranska M, Schulz H, Simon PW, Nothnagel T (2006) Biopolymers 81:497
92. Baranska M, Schulz H (2006) Planta 224:1028
93. Baranska M, Schulz H, Christensen LPJ (2006) J Agric Food Chem 54:3629
94. Baranska M, Schütze W, Schulz H (2006) Anal Chem 78:8456
95. Baranska M, Baranski R, Grzebelus E, Roman M (2011) Vib Spectr 56:166
96. Brackmann C, Bengtsson A, Alminger ML, Svanberg U, Enejder A (2011) J Raman Spectr 42:586
97. de Oliveira VE, Castro HV, Edwards HGM, de Oliveira LFC (2010) J Raman Spectrosc 41:642
98. López-Sánchez M, Ayora-Cañada MJ, Molina-Díaz A (2010) J Agric Food Chem 58:82
99. Ivleva NP, Niessner R, Panne U (2005) Anal Bioanal Chem 381:261
100. Schulte F, Lingott J, Panne U, Kneipp J (2008) Anal Chem 80:9551
101. Schulte F, Mäder J, Kroh LW, Panne U, Kneipp J (2009) Anal Chem 81:8426
102. Sanchez-Cortes S, Domingo C, Garcia-Ramos JV, Aznarez JA (2001) Langmuir 17:1157
103. Jones H, Safe S, Thaller V (1966) J Chem Soc C 52:1220
104. Bohlmann F, Burkhardt T, Zdero C (1973) Naturally occurring acetylenes. Academic Press, London
105. Xu X, Yuan Y, Wang X, Li Y (2011) FEMS Microbiol Lett 314:42
106. Minto RE, Blacklock BJ (2008) Prog Lipid Res 47:233
107. Even J, Bertault M, Girard A, Delugeard Y, Fave J-L (1994) Chem Phys 188:235
108. Schrader B, Schulz H, Baranska M, Andreev GN, Lehner C, Sawatzki J (2005) Spectrochim Acta A 61:1395
109. Roman M, Dobrowolski J Cz, Baranska M (2011) J Chem Inf Model 51:283
110. Baranska M, Schulz H, Baranski R, Nothnagel T, Christensen LP (2005) J Agric Food Chem 53:6565
111. Roman M, Dobrowolski J Cz, Baranska M, Baranski R (2011) J Nat Prod 74:1757

112. Metzger BT, Barnes DMJ (2009) Agric Food Chem 57:11134
113. Frese L (1983) Gartenbauwissenschaft 48:259
114. Roman M, Baranski R, Baranska M (2011) J Agric Food Chem 59:7647
115. Alvarez L, Marquina S, Villarreal ML, Alonso D, Aranda E, Delgado G (1996) Planta Med 62:355
116. Redl K, Breu W, Davis B, Bauer R (1994) Planta Med 60:58
117. Chang M-H, Wang G-J, Kuo Y-H, Lee C-K (2000) Chin Chem Soc 47:1131
118. Thygesen LG, Jørgensen K, Møller BL, Engelsen SB (2004) Appl Spectrosc 58:212
119. Micklander E, Brimer L, Engelsen SB (2002) Appl Spectrosc, 56:1139
120. Nersissian AM, Mehrabian ZA, Nalbandyan RM, Hart PJ, Fraczkiewicz G, Czernuszewicz RS, Bender CJ, Peisach J, Herrmann RG, Valentine J (1996) Protein Sci 5:2184
121. Nestor L, Larrabee JA, Woolery G, Reinhammer B, Spiro TG (1984) Biochemistry 23:1084
122. Fălămaş A, Pînzaru SC, Dehelean CA, Peev CI, Soica C (2011) J Raman Spectrosc 42:97
123. Talian I, Oriňák A, Efremov EV, Ariese F, Kaniansky D, Oriňáková R, Hübner J (2010) J Raman Spectrosc 41:964
124. Baranska M, Schulz H, Siuda R, Strehle MA, Rösch P, Popp J, Joubert E, Manley M (2005) Biopolymers 77:1

Chapter 10
Chemometric Analysis of Raman and IR Spectra of Natural Dyes

Anna Rygula and Paweł Miskowiec

Abstract The use of chemometric methods in the analysis of Raman and IR spectra of dyes has opened new possibilities in the areas previously reserved only for chromatographic methods or conventional UV-vis studies. This chapter is focused on the analysis of natural dyes, especially carotenoids, flavonoids, anthraquinones and indigoids presented in biological samples e.g. plants or food. The discussed chemometric methods are implemented both in the qualitative and quantitative determination of colorants with the techniques of vibrational spectroscopy

Keywords Natural dyes • Antraquinones • Flavonoids • Indigoids • Carotenoids • FTIR • NIR • Raman spectroscopy

10.1 Introduction

Chemical and physical analysis of dyes is an important aspect of the biological and medical investigations. Firstly, it benefits with a fundamental knowledge about live organisms and properties of natural chromophores. The next point is the usage of natural dyes in food, energy and textile industry and the necessity of determination them in many consumer products. The last but not least, is an application of dye analysis in archaeology and conservation of art. It should be noted that this chapter will be focused mainly on the dye analysis in natural samples e.g. plant or lichens and their extracts.

Generally speaking for the detection and qualitative/quantitative determination of bioactive compounds, chromatographic methods have been routinely used [1]. Despite of their popularity, the problem is connected with the pretreatment of the sample prior to its analysis. These techniques require to extract the analyte from the tissue matrix what can lead to loss the information about its interactions with matrix and even to pigment degradation. In this context, the possibility of non-destructive

A. Rygula (✉) · P. Miskowiec
Faculty of Chemistry, Jagiellonian University, Ingardena 3,
30-060, Krakow, Poland
e-mail: barana@chemia.uj.edu.pl

M. Baranska (ed.), *Optical Spectroscopy and Computational Methods in Biology and Medicine,* Challenges and Advances in Computational Chemistry and Physics 14,
DOI 10.1007/978-94-007-7832-0_10, © Springer Science+Business Media Dordrecht 2014

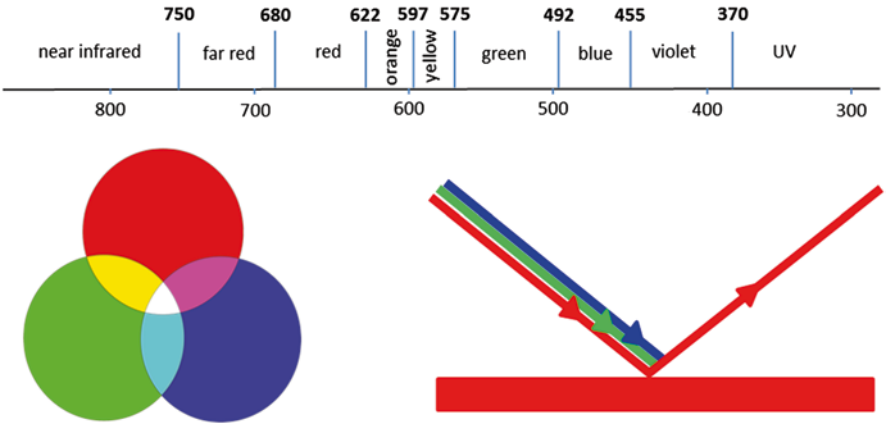


Fig. 10.1 The spectral range of visible light (based on [7]). Below *left*, the scheme of primary colours which compose a white light and below *right* the principle of colour observation for objects

or *in situ* analysis with the vibrational spectroscopy is attractive and will be presented in the following chapter.

According to the definition [2, 3], colorants are classified as pigments i.e. small particles, practically insoluble in those media in which they are applied, and dyes which are soluble in the solvent. Talking about the natural colorants in biological tissues, these definitions are usually not applied and both terms are used alternatively. It is usually caused by limitation of knowledge about form of dye occurrence.

Colorants are characterized by their ability to absorb or emit light in the visible range (400–700 nm) (Fig. 10.1) [4]. The colour of matter is visible when the white light, composed of all wavelengths from the visible spectrum, is partially absorbed whilst the reflected part of light is observed. Generally, organic dyes possess colour because they have at least one chromophore (a colour-bearing group). The most common are the molecular structure of dye including a conjugated system responsible for the resonance of electrons [5, 6].

10.1.1 The Classification and Function of Dyes in the Natural World

Colours in nature are so popular that we usually do not think about them and do not realize how important they are in our life. The main source of natural pigments are plants i.e. leaves, fruits and flowers. The animal world is also full of colour as pigments are present in skin, eyes, and other tissues but also in bacteria and fungi.

The role of pigments in the nature is difficult to overestimate but at the same time, our knowledge about their distribution, availability and properties is still limited [8]. The coloration patterns of flowers attract insects and contribute to the pollinating

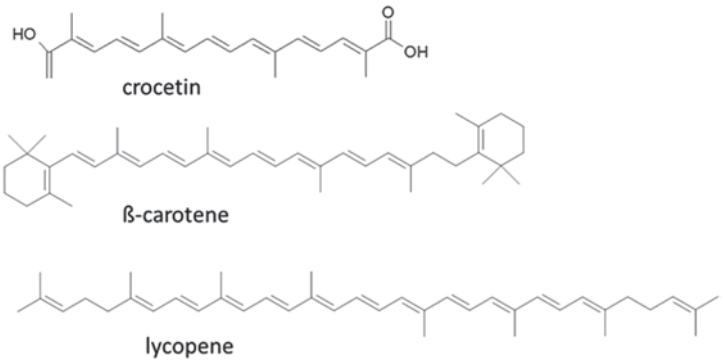


Fig. 10.2 Examples of carotenoids

process and it is well known that many species accumulate colorants e.g. flavonoids in the reproductive organs. Likewise animals are attracted by fruit colour, and after the consumption the seeds are dispersed through the excrement [4].

The concentration of natural dyes in plants is generally low and it varies depending on species and pigment. For anthocyanins these amounts are about 0.1–5 g/100 g dry matter [9], whilst for lycopene and β-carotene the mean levels are between 4.5–6.3 and 0.8–1.6 mg/100 g of fresh matter, respectively [10]. What is important, a lot of dyes, especially flavonoids and anthraquinones, occurs in plant tissues in the form of glycosides, where molecule of dye is bound to carbohydrate moiety e.g. glucose.

The natural pigments can be divided according to the structural characteristics and here this classification will be applied [8]. Only the most important pigment groups in regard of application of vibrational spectroscopy techniques are discussed below.

Carotenoids (Fig. 10.2) are usually built from eight isoprenoid units and belong to the class of hydrocarbons (carotenes) or their oxygenated derivatives (xanthophylls). The most characteristic part of them is long central chain composed of a conjugated double-bond system which acts as a chromophore and is responsible for yellow, orange, or red colour of these compounds [11].

They have been identified in photosynthetic and nonphotosynthetic organisms. Carotenoids are responsible for many of the brilliant red, orange, and yellow colours of fruits, vegetables, fungi and also of birds, insects, crustaceans and fishes [8]. It is interesting that only plants and bacteria can synthesize carotenoids *de novo*. Particularly many carotenoids have been isolated from marine organisms. Carotenoids are also accumulated in chloroplasts of all green plants as a mixture of α- and β-carotene, β-cryptoxanthin, lutein, zeaxanthin, violaxanthin and neoxanthin [6, 8].

Carotenoids are the main pigments, apart from chlorophylls, involved in photosynthesis. They have two well-known functions: as (1) accessory pigments in light

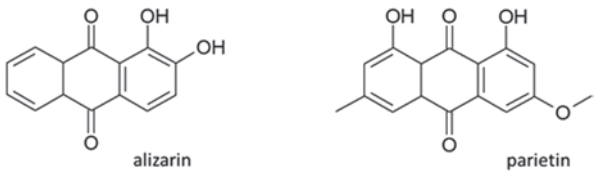


Fig. 10.3 Examples of anthraquinones

harvesting and (2) photoprotectors against oxidative damages [12]. It has been established that their antioxidant activity and the ability of quenching of singlet oxygen is related to the structure, precisely to conjugated double-bond system. Many diseases, such as cancer and strokes, involve oxidative processes mediated by free radicals. Carotenoids, by their antioxidant effect, can show benefits in such diseases; however, this function is not completely demonstrated *in vivo* [11].

In conjugated polyenes group, one can separate the subgroup of psittacofulvins or parrodienes. These pigments are widely distributed in corals mollusc, pearls and parrots' plumage [13–15].

Quinones are more widely distributed than the other natural pigments (with the exception of carotenoids and melanins) [6]. According to their structure, they can be divided into benzoquinones, naphthoquinones and anthraquinones. One of the most popular benzoquinoid dyes is carthamine, whilst the representative examples for naphtoquinones are juglon and lawson received from green walnut shell (Juglandaceae family) and henna (*Lawsonia inermis* L.), respectively. The biggest group of quinones is formed by anthraquinones (Fig. 10.3) where one of the most popular anthraquinoid dyes are alizarin and purpurin, traditionally obtained from madder (Rubiaceae family) [16] and cochenille. The another famous anthraquinoid dyes are carminic and laccaic acids, which have been traditionally obtained from insects *Dactylopius coccus* L. and *Laccifer lacca*, respectively.

Many quinones are by-products of the metabolic pathways and only a few organisms (fungi) produce them in large quantities. In general, quinones give yellow, red or brown colorations but quinone salts reveal purple, blue or green colours [6].

Flavonoids are colorants responsible for the colour of many food products, especially fruits and vegetables [17]. Moreover, this is the most widely distributed group of secondary metabolites consumed by people and at the same time, the most studied. More than 6000 flavonoids have been already discovered and isolated from plants [6, 17, 18]. Flavonoidal colorants show a wide distribution in vascular plants. These compounds give colours from yellow and orange to blue for such parts of plants as petals, fruits, leaves and roots [8].

Flavonoids are derivatives of a 2-phenylchromone structure where three phenolic rings are described as the A-, B-, and C-rings. Depending on substitute of the functional group and the modification to the A, B, and C rings, flavonoids are divided into several classes [17, 19] (Fig. 10.4).

Flavonoids selectively absorb light of blue-green and red wavelengths, which is important for the photosynthetic process and the formation of patterns on petals pigments. At the same time they block ultraviolet radiation, which is destructive to

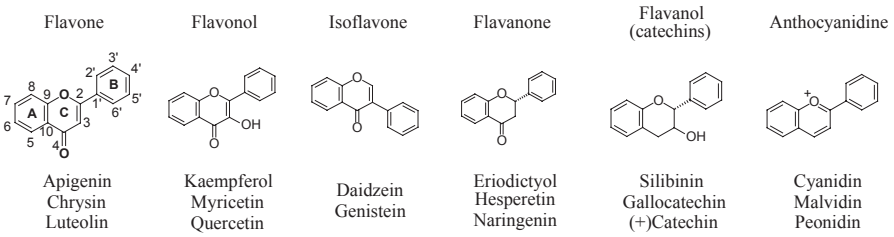


Fig. 10.4 The classes of flavonoids, their structures and examples of representatives

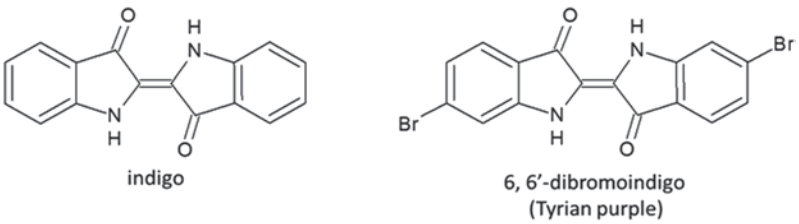


Fig. 10.5 Indigoids

nucleic acids and proteins [20, 21]. Moreover, it has been clearly established that flavonoids in natural products (e.g. grape, soybean, peanut, wine, tea) have a good antioxidant activity and pharmacological effects [21, 22 and refs cited therein].

Due to the ability of forming different chemical structures depending on pH, anthocyanins are one of the most frequently studied class of flavonoids. The most stable form exists at acidic pH level (pH = 1) and is known as the oxonium or flavylium cation, giving bright orange–red to violet colour. In neutral pH (pH = 7), anthocyanin gives blue–green colour and the most abundant form is quinoidal base. In pH = 4.5 anthocyanins exist in the chalcone and/or hemiketal forms which are colourless [17]. Due to this acid-base equilibrium, the most of red and blue colours occurred in plants are produced by water-soluble anthocyanin pigments, which are located in vacuoles [23].

The other group of dyes, indigoids, are described as a bisindole alkaloids. The most famous representatives are indigo, indirubin, and 6, 6'-dibromoindigo (Tyrian purple) (Fig. 10.5).

The natural sources of indigo were widely known since centuries. In moderate climate the indigo dye was produced from woad (*Isatis tinctoria*), which is native for Europe and dyer's knotweed (*Persicaria tinctoria*, formerly known as *Polygonum tinctorium* [24]) used both in Europe and Asia (Japan). In tropical climate, this dye was obtained from the genus *Indigofera* and it was probably the most popular source of indigo before developing of indigo synthesis in the 90's of nineteenth century. Apart from indigo, the natural received pigment contains some amount of indigo derivatives where the most significant is indirubin (about 2–15%) [25].

Very famous indigoid dye is 6, 6'-dibromoindigo also known as Tyrian or Royal purple. It is one of the oldest known and the most expensive dye. The colour is obtained from marine shellfish of the Muricidae and Thasidae families [26].

10.2 Techniques Used for Dye Analysis

10.2.1 Vibrational Spectroscopy Techniques

Raman spectroscopy can obey the restrictions and terms related to the analysis of natural dyes and pigments. The main problem is very low amount of the analyte in a sample and in consequence, the necessity of applying very sensitive techniques. The second difficulty is the absorption of the visible light, which is a characteristic feature of the chromophores. This phenomenon is responsible for the fluorescence, which perturbs and overlaps the weak Raman signal. Moreover, the fluorescence of the remains pollutants of the biological components of the sample is also possible. To avoid the latter, the near infrared wavelength lasers are applied (FT techniques) [11, 27, 28].

The absorption is responsible also for the resonance or pre-resonance effect applied in the Resonance Raman spectroscopy (RRS). It is observed when the incident laser light occurs in the range of an electronic absorption band of the analysed compound. As a consequence, the Raman signal of absorbing compound is enhanced by the factor of 10^6 [29] that allows to detect the analyte despite its very low concentration. Taking into account that the most popular lasers provide an excitation in the visible or NIR range, it is clear that this effect could be observed frequently in the dye analysis.

Several advantages in dye analysis provides the Surface Enhance Raman Spectroscopy (SERS). This method requires an interaction of the sample with metal nanoparticles, usually the silver colloid. According to the theory [29, 30], this interaction causes the chemical and electromagnetic enhancement of the Raman signal up to 10^{15} times and consequently, an intensification of the Raman spectra. SERS is a popular method in a dye analysis especially in the field of art conservation. There is a few reports with the application of this technique for the analysis of natural dyes in textiles [31–34]. The use of SERS to the *in situ* analyses of natural samples is still discussed and developed [35].

To sum up, it is worth to underline that during analysis of dyes it is a good practice to work with different wavelengths of lasers both for avoiding of the fluorescence and taking the advantage of the resonance Raman effect. However the techniques described above are useless when the Raman signals of pigments are too weak and/or are overlapped by other biological components of sample. In this situation, the chemometric methods can significantly expand the scope of applicability of the method.

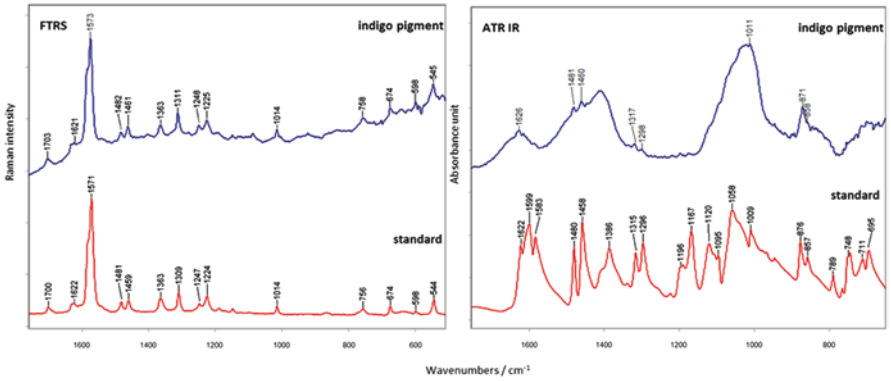


Fig. 10.6 FT-Raman (*left*) and ATR IR (*right*) spectra of natural indigo pigment (*Isatis tinctoria*) compared to synthetic standard of indigo

Infrared spectroscopy can be also used for the analysis due to the fact that molecules absorb infrared light of specific frequencies. One can divide the IR range into: near infrared (14,000–4,000 cm^{-1}), mid infrared (4,000–400 cm^{-1}) and far infrared (400–10 cm^{-1}). For dyes, the near and mid infrared spectroscopies have been exploited for decades [36, 37] in biology, chemistry, forensic science, cultural heritage and many more. Those methods owe the popularity mainly due to their rapidity of the analysis and low consumption of chemicals.

In mid-infrared dye analysis the reflectance methods are principal. Amongst them ATR (attenuated total reflection) technique is commonly used due to its capability to measure *in situ* and no need for sample preparation [38]. Apart from ATR, a DRIFT technique (Diffuse Reflectance Infrared Fourier Transform Spectroscopy) is exploited for qualitative analysis of some dyes [39, 40] even if the dye structure is unknown. However, DRIFT technique is particularly useful in the investigations of the powdered samples. That is why it is more under the subject of interests of industry and conservation sciences than biology or medicine.

Both techniques—RS and IR—are considered as complementary methods and a good example of that is the analysis of indigo pigment [41]. In chosen work the analyses were focused on the detection of the pigment in the textile samples and comparing its natural and synthetic sources. It was shown that Raman spectroscopy is a very sensitive method to identify indigo both on the colouring thread and in the presence of sample impurities. The Raman signal of the pigment is so strong that any other bands are almost invisible (Fig. 10.6). At the same time, ATR IR technique can provide information about the remaining components which are presented in the sample e.g. impurities. Simultaneously the indigo bands are not so clearly observed. Generally, the sensitivity of IR technique for indigo detection is not so high as of FT-RS. Nevertheless, IR technique can provide information about the source and complexity of indigo pigments, in contrast to FT-RS which is better for detection even the traces of this pigment.

The growing interest in using of IR and Raman spectroscopies instead of commonly applied chromatographic and UV-vis methods to quantitative and qualitative analysis of dyes, has been dictated by a number of factors. The most important of them are rapidity, simplicity, safety to the environment (extremely low chemicals consumption) and low operational costs [42–45]. In most cases infrared methods can be also applied in the non-destructive way [45, 46]. The mid-IR spectroscopy has been successfully used to differentiate the color ink and paint entries [44–47]. However, application of mid-IR for dyes analysis in forensic science as well as in cultural heritage studies is beyond the scope of this chapter.

The assistance of chemometrics to mid-IR spectroscopy is often related with the evaluation of the results against HPLC and even UV-vis data. The latter are typically used in dyes determination and have been developed as reference methods. Mid-IR spectroscopy is also used in the food quality control. The results are in most cases comparable to those obtained with traditional UV-vis and HPLC methods [48–50].

10.2.2 Chemometric Methods Used for Quality, Quantity and Distribution Analysis

According to [51], two chemometric approaches can be named: uni- (or bivariate) and multivariate analysis. Uni- or bivariate analysis uses one or two variables to extract information from a system e.g. peak area or peak height. This approach can be difficult to use when many peaks overlap but is the easiest and the most intuitive. It is the most popular approach to analyse mapping results.

Multivariate analysis uses many (up to thousands) of variables to obtain quantitative data. Very small, even non-distinguishable by human eye spectral changes can give significant meaning and provide important results. The main advantage is the possibility to analyse the overlapping peaks and simultaneous detection of several components in a sample.

Figure 10.7 shows the classification of the major methods used in analysis of spectroscopic data [52]. Presented in Fig. 10.7 graph was dedicated to NIR spectroscopy, but the same rules could be applied to other spectroscopic techniques.

In recent years, NIR spectroscopy combined with chemometrics has attracted considerable attention in chemical content analysis. The main advantage of NIR comparing to mid-IR spectroscopy is that using NIR spectroscopy one can typically penetrate sample much deeper. NIR spectroscopy can be therefore, very useful in probing bulk material with very little or even no sample preparation.

Most of absorption bands in the NIR range are overtones or combination of the fundamental vibrations of the molecules. The typical NIR reflectance spectra and its first derivative are shown in Fig. 10.8 [53]. They were obtained for banana pulps with high content of carotenoids, prepared from different genome groups. One can observe that in the complex mixtures, such as biological material, food or medicines, the multiple bands and the effect of peak-broadening result in broad signal

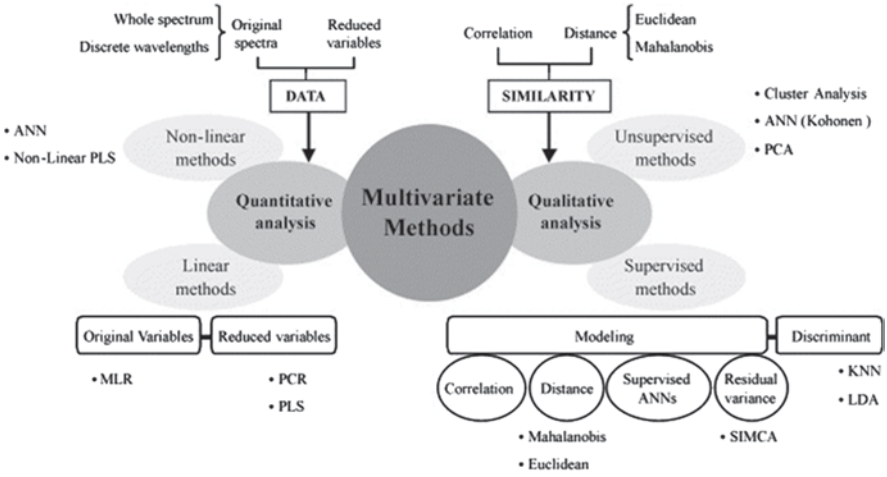


Fig. 10.7 Classification of the major qualitative and quantitative multivariate-analysis methods used in NIR spectroscopy. *ANN* Artificial Neural Networks, *PCA* Principal Component Analysis, *PLS* Partial Least Squares, *PCR* Principal Component Regression, *MLR* Multiple Linear Regression, *SIMCA* Soft Independent Modeling Class Analogy, *KNN* K-Nearest Neighbour, *LDA* Linear Discriminant Analysis. (Reproduced with permission from Ref. [52]. © (Elsevier) (2013))

shapes with few sharp peaks. For other typical NIR spectra of various color fruits see [54].

Obviously, the analytical information contained in the broad, extensively overlapped bands of NIR spectra is hardly selective and multivariate in nature. In addition, differences between samples may cause very slight spectral differences that are difficult to distinguish. This similarity is the reason why the sophisticated multivariate chemometrics techniques are essential to extract useful information from NIR spectrum and to perform qualitative or quantitative NIR analysis. To extract analytically important and reduce irrelevant information the following pretreatment mathematical methods are used practically in every NIR analysis, including dyes:

- Multiplicative scatter correction (MSC) and standard normal variate (SNV)
- Reduction of variables by principal component analysis

To do any further quantitative analysis the spectrometer has to be calibrated using multivariate methods, which are similar to those used in mid-IR spectroscopy. The most frequently multivariate regression methods applied in quantitative NIR analysis are principal component regression and partial least-squares regression. Qualitative multivariate analytical methods are known as pattern-recognition methods, and are subdivided in “supervised” and “non-supervised” depending on whether or not the class to which the samples belong is known [55]. Some other information of both qualitative and quantitative chemometrics methods, are described in subchapter 3. More detailed description the reader can find elsewhere [55, 56].

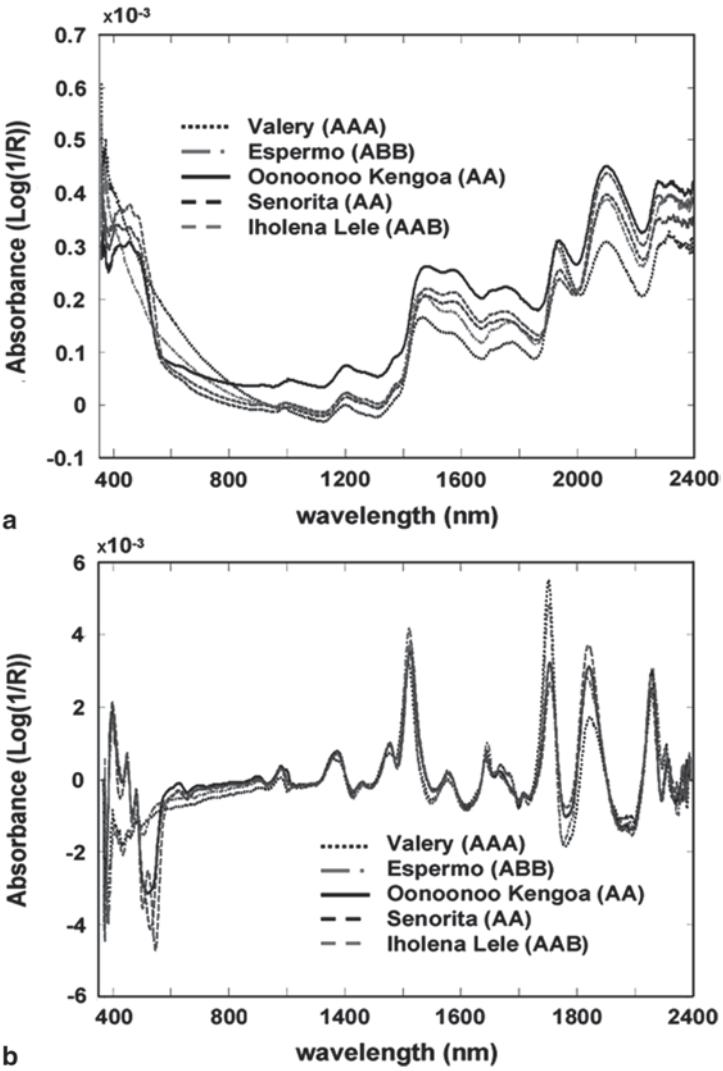


Fig. 10.8 Typical Vis/NIR spectra for *Musa* fruit pulp samples from varieties of different genome groups (**a**) without preprocessing and (**b**) Savitzky-Golay smoothed first derivative of the spectra, as described in [53]. (Reproduced with permission from Ref. [53]. © (ACS Publications) (2013))

Taking into account these advantages, the NIR has been applied successfully in dye analysis [53, 57–59]. Multivariate calibration methods such as multiple linear regression (MLR), principal component regression (PCR) and partial least squares (PLS) are typically and successfully used to determine the chemical composition of analyzed dyes. One can even state, that PLS is a method that can be regarded as a typical tool in chemometric analysis of the NIR data obtained from dyes.

Apart from single *in situ* measurements the mapping of biological tissues can be performed followed by the chemometric analysis [60]. First, the spectra are collected from the surface of sample and then data are processed to get the chemical image of the sample. The most popular method is the integration of the spectral intensity signal around the spectral position of characteristic vibrational bands of the pigments but also hierarchical cluster analysis (HCA) is often used. Finally the 2D or 3D chemical image of the distribution and relative concentration of the analyte is obtained. It should be remembered that this technique often requires to define the so called marker band or characteristic spectral range of substances under study, what is very often problematic. The Raman (or IR) bands may be useful as markers for Raman (or IR) sample mapping, as long as one can avoid band overlap hindering a clear separation of other biological components e.g. lignin, cellulose, chlorophyll [9, 27, 28, 61] or proteins [9].

In case of the problem with defining of marker bands, also the chemometry techniques could be used. Only the few examples are briefly described below.

Principal component analysis (PCA) is often used for spectroscopic analysis. The main idea of that method is an extraction of data and in consequence, reduction of the number of variables that contain maximal variance. In some cases, the number of different components could be determined with this method [62]. The similar application has linear discriminant analysis (LDA) which is also applied to data classification.

Cluster analysis (CA) groups data according to their similarities. The most popular types of CA are hierarchical CA (HCA) and K-means CA. Finally, one gets a dendrogram, where the dependence between data is represented by a branches of tree or, in case of imaging techniques, certain number of areas which shows a distribution of various clusters in the surface [62, 63].

The interesting method of computational analysis of spectra is 2D-correlation spectroscopy, which originates from 2D NMR spectroscopy and was developed by Isao Noda in 80's. The idea of the 2D correlation analysis is to subject a sample to an external perturbation while all other parameters of the system are kept at the same value. Spectra of the initial sample and subjected to the perturbation are correlated one to another. The nature of perturbation can be different e.g. temperature, pressure, pH or changes of the chemical composition of the system. That computational method could be a convenient tool for detection and analysis even minor changes of the spectra. More information about this method one can find elsewhere [64].

The background for quantitative analysis is the fact that Raman scattering is directly proportional to the concentration of the vibrating species [51]. In case of IR techniques, the absorption is proportional to the concentration of the analyte according to the Lambert-Beer law. Both methods can be successfully supported in quantification analysis with the use of chemometric methodes.

In case of IR spectroscopy chemometric support has been widely used mainly in near infrared range, especially during *in situ* measurements. This is because NIR measurements might be influenced by a high water content of a sample, which strongly disturbs the analysis, especially from the water bands which are additionally enhanced by their multiplication (overtones) [10].

Table 10.1 The wavenumbers positions of ν^1 vs. the number of conjugated double bonds in the carotenoid chain. (adapted from [11])

Carotenoid	Number of conjugated double bonds in the carotenoid chain	Wavenumber [cm^{-1}]
Crocetin	7	1,536
β-carotene	9	1,524
Lycopene	11	1,510

On the other hand, it should be noted that in both Raman and IR studies of the natural products (including dyes) the quantitative comparison with other, the so called reference techniques is unavoidable. In such operation chemometrics is also irreplaceable.

The most popular techniques used for quantitative analysis are classical least square (CLS), multiple linear regression (MLR), partial least square (PLS) and principal components regression (PCR) [51, 65].

The PLS method applied in RS is so-called "full spectrum method" what means that calibration model is improved with an increasing number of data points, in opposite to methods which require to choose the marker bands [66].

10.3 Raman and IR Analysis of Selected Dyes

In this chapter more details are provided for carotenoids, flavonoids, antrachinons and indigo, in addition to information gathered to illustrate the qualitative and quantitative FT-Raman analysis of plants (Chap. 9). Moreover in this part of review Raman data are followed by IR and NIR results for these dyes.

10.3.1 Carotenoids

The Raman analysis of carotenoids is well documented [e.g. 11, 67]. Carotenoids occur in plants as a secondary metabolites on the ppm level, however, due to high sensitivity of the Resonance Raman spectroscopy (with the laser excitation from the visible range), even such small amounts are possible to be detected [68, 69]. A strong enhancement can be also achieved by using FT-Raman spectroscopy and pre-resonance effect.

Moreover, it is known that the wavenumber location of the marker carotenoid bands, especially C=C, is correlated with the length of the polyene chain (Table 10.1), i.e. the position is red-shifted with the extent of the conjugation length of the polyene chain due to an electron-phonon coupling. Moreover, the band position could be slightly influenced by the carotenoid side group and bonding to other plant constituents. It is explained by the fact that the C=C stretching wavenumber of carotenoids is sensitive to the complexity of the matrix in which this compound is involved [11, 70].

One of the most noteworthy example of carotenoid mapping in the biological tissue is reported by S. Gamsjaeger et al. [71]. The study presents the *in situ* discrimination of carotenoid pigments (carotens and xanthophylls) beside the flavonoid substances (anthocyanins and flavonols) in petals of *Viola x wittrockiana* (Fig. 10.9). The spectra were collected with laser with wavelength at 1,064 nm.

These results confirm that chemical differences in flower petals correlated to the molecular structure of pigments can be directly detected by Raman spectroscopy combines with HCA. It is worth to notice that in this case, the whole "fingerprint" spectral range (1,000–1,700 cm^{-1}) was analyzed which includes all Raman bands from the flower matrix e. g. lignin and cellulose.

The three lowest rows of Fig. 10.9 display the distribution of carotenoids, anthocyanins and flavonols obtained from the integrate band intensity within the spectral marker regions individually defined for the every pigment. This technique gives the information about the relative pigment concentrations. The map of carotene distribution was created by using the band around 1,156 cm^{-1} because it shows little or no overlap with Raman bands ascribed to flavonol or anthocyanin derivatives, which were analyzed simultaneously.

The more detailed FT-RS study of carotenoids was conducted by Schulz et al. [11] for *Calendula officinalis* L. The distribution of carotenoids with different lengths of conjugated double-bond chain: auroxanthin (7), luteoxanthin (8) lutein (9) and antheraxanthin (9), was shown (Fig. 10.10).

The carotenoids were analyzed also in the algal cell. By using the FT-Raman microimaging, the astaxanthin (classified as xanthophyll) was *in situ* and *in vivo* detected in *Haematococcus pluvialis* [72]. It was revealed that the structure of astaxanthin in algal cell is different from the structure of synthetic compound, what was explained by the change of the chemical environment of the carotenoid chain. The Raman maps of carotenoid distribution in the algal species *Neochloris oleobundans* was reported in ref. [73]. The carotenoid compound was detected simultaneously with lipids (triglyceride).

The resonance Raman effect is applied in the measurement of macular carotenoids in the human eye. The medical studies suggest that macular pigment, lutein and zeaxanthin, protects against visual loss from age-related macular degeneration, the leading cause of irreversible blindness [74]. The analyses can be carried out *in vivo*, with the green or blue lasers [75–77]. The similar studies on Raman detection of carotenoids were conducted also on human blood [78] and skin [79].

In regard of similarity in chemical structure between carotenoids and parrodienes, the Raman analysis of these pigments has a similar requirements. Polyenes are detected, usually accompanied by carotenoids, i.e. in corals [13, 80] and molluscan shells [81, 82]. The two major bands are observed at ca. 1,500 cm^{-1} (v C=C) and 1,130 cm^{-1} (v C-C) for all specimens [80].

Raman spectroscopy can be used for *in situ* quantitative measurements of carotenoids. The analysis of crocetin in saffron (*Crocus sativus* L.) was carried out with applying of RS and chemometrics [83]. A rapid method for determining the crocetin esters and colouring strength directly in saffron samples, using NIR Raman spectroscopy (785 nm) together with partial least-square regression (PLS) was

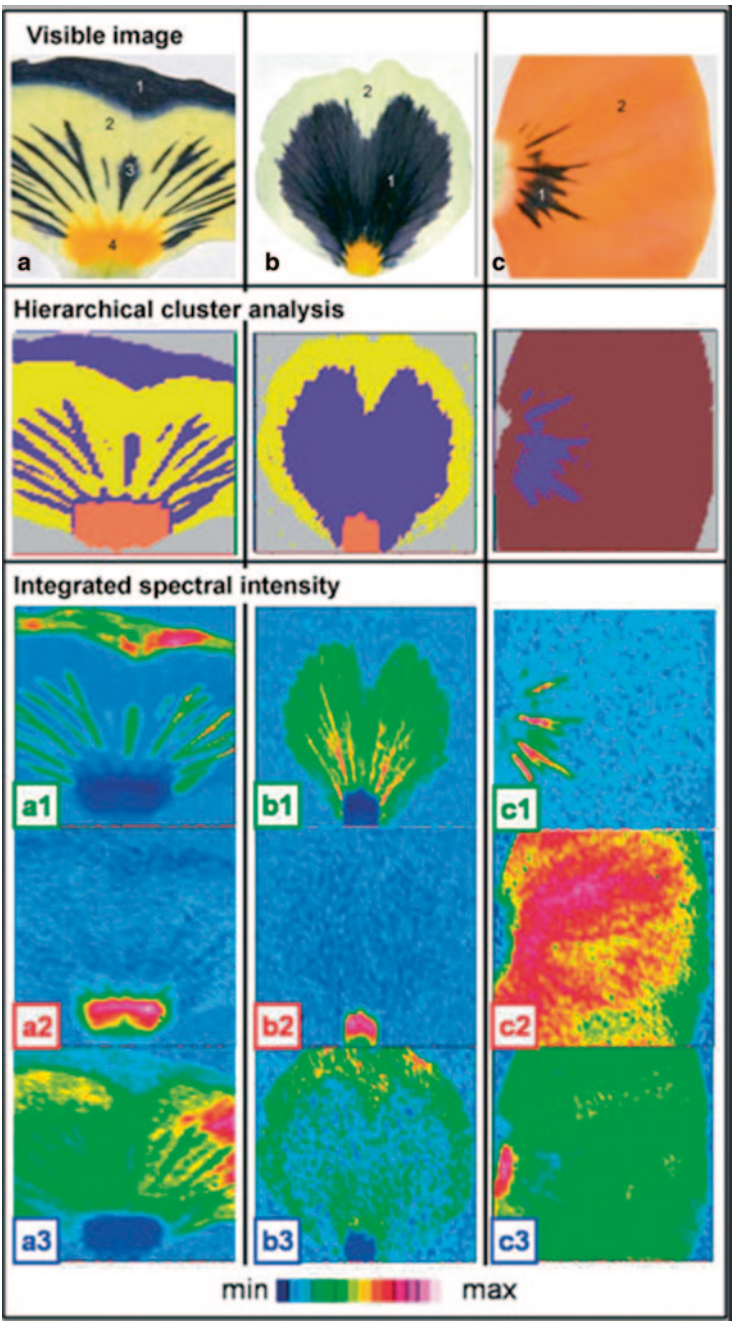


Fig. 10.9 The first two rows show the visual images (a, b, and c) of petals of *Viola x wittrockiana* and the corresponding chemical images after HCA, obtained from two-dimensional Raman maps (a: mapping area14 × 22.75 mm, increment 250 μm; b: mapping area 34.58 × 22.42 mm, increment 380 μm; c: mapping area 20.75 × 15.75 mm, increment 250 μm). The colours in the chemical image

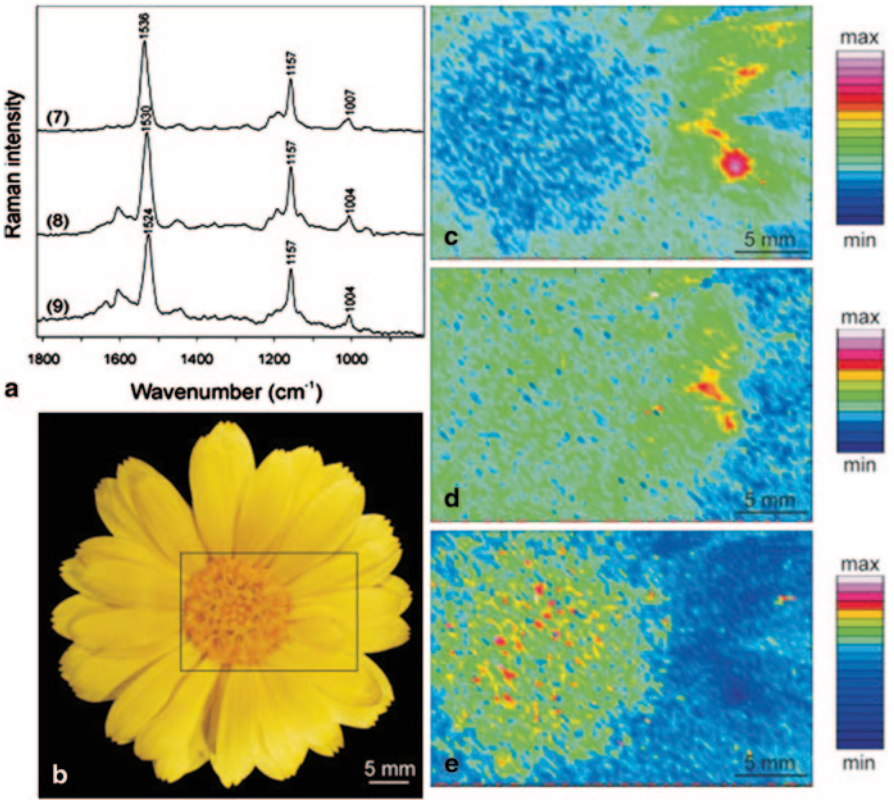


Fig. 10.10 FT-Raman spectra of *Calendula officinalis* L. measured in three different points showing the presence of 7-, 8-, and 9-conjugated carotenoids (**a**). Picture of *Calendula officinalis* L. flower (**b**) and corresponding Raman maps coloured according to the band intensity at 1,536 (**c**), 1,530 (**d**), and 1,524 cm^{-1} (**e**) related to the content of 7-, 8-, and 9-conjugated double bond carotenoids, respectively. (Reproduced with permission from Ref. [11]. © (Wiley) (2013))

demonstrated [83]. As a reference methods HPLC and UV-vis spectroscopy were used. The model in the spectral region 1,700 cm^{-1}–955 cm^{-1} showed the best correlation coefficient. According to the authors, the models can be used as a tool for rapid screening of quality in saffron samples. Moreover, RS gives better results and is more adequate than NIR spectroscopy for the study of saffron carotenoids.

---

show clusters of Raman spectra with the least heterogeneity. HCA according to Ward's algorithm in the wavenumber range 1,000–1,700 cm^{-1}, with vector normalization as data pre-processing. The chemical images in the last three rows show the patterns resulting from the integrated spectral intensity: in a1, b1, and c1 for the anthocyanin marker (integration range 1,225–1,272 cm^{-1}), in a2, b2, and c2 for the carotenoid marker (integration range 1,140–1,172 cm^{-1}), in a3, b3, and c3 for flavonol glycoside marker (integration range 1,550–1,588 cm^{-1}). The colour scale correlates with the magnitude of the integrated intensity and is a signature of the relative concentration. (Reproduced with permission from Ref. [71]. © (Wiley) (2013))

The analysis of lycopene and β-carotene in tomatoes fruits and their products using FT-Raman, ATR IR and NIR spectroscopy gives similar results [10]. The best prediction was achieved for PLS model based on the ATR IR spectroscopy (R^2 = 0.98 and −0.97, SECV = 33.20 and 0.16 for lycopene and β-carotene, respectively), whilst the worst prediction was shown for NIR spectroscopy (R^2 = 0.85 and 0.80, SECV = 91.19 and 0.41, respectively) [8]. It may be caused by a high water content in all tomato samples which dominate in the NIR spectra. It is important to notice, that the prediction quality for ATR IR spectroscopy in the 650–1,800 cm^{-1} is better than for FTRS, even when the wavenumber range was limited to marker carotenoid band. However, FT-Raman (1,064 nm) spectroscopy enabled to discriminate lycopene (1510 cm^{-1}) and β-carotene (1,520 cm^{-1}) signals what is impossible using laser excitation in the visible range [84].

The β-carotene and parietin (antraquinon) were analyzed in Antarctic lichens using FT-RS combined with cluster analysis (CA) and principal component analysis (PCA) [85]. The aim of that study was to obtain relative quantitative information about changes of pigment concentration in samples.

The interesting example of a carotenoid detection in the industrial approach is the determination of the total carotenoid concentration in dried pastas [86]. The authors compared the different techniques like HPLC, spectrophotometry, colorimetric and photoacoustic spectroscopy (PAS) with the resonance Raman technique. The results obtained for RS was satisfactory however, the PAS was found as the most favourable method mainly because of better sensitivity and repeatability of measurements.

The RS could be potentially applied to the analysis of carotenoids in fruit and vegetable juices [84]. The analysis of grape, orange, tomato, carrot and vegetable mix products were carried out using 488 nm laser and the height of band at 1,525 cm^{-1} was quantified for data processing. It is worth notice that better correlation was obtained for diluted samples (R = 0.98, P = 0.003) than for undiluted (R = 0.94, P = 0.01).

The potential of mid-infrared spectroscopy for the chemical study of carotenoids, have been established a few decades ago [87]. However, the recent research has depicted the capability of this technique for the simultaneous quantitative and qualitative determination of carotenoids i.e. lycopene and β-carotene in fruits like tomatoes [43, 88]. Fig. 10.11 depicts an example of the correlation between lycopene content in different tomato varieties measured by mid-IR and HPLC methods.

The authors analyzed the 950–980 cm^{-1} infrared region where C = C trans double bonds absorb and used partial least square regression (PLSR) to produce calibration models that correlate lycopene or β-carotene concentration in standards (as determined by UV-vis spectroscopy) with their infrared spectra. The achieved correlation factors were satisfying (R^2 = 0.95), thus the PLSR calibration models were then used to predict the lycopene concentration in commercial and experimental tomato varieties using their infrared spectra.

Those research were coherent with earlier studies [10] showing the application of three different types of spectroscopies (FT-Raman, ATR IR, and NIR Spectroscopy) in carotenoids analysis. As a results authors got the best correlation with HPLC

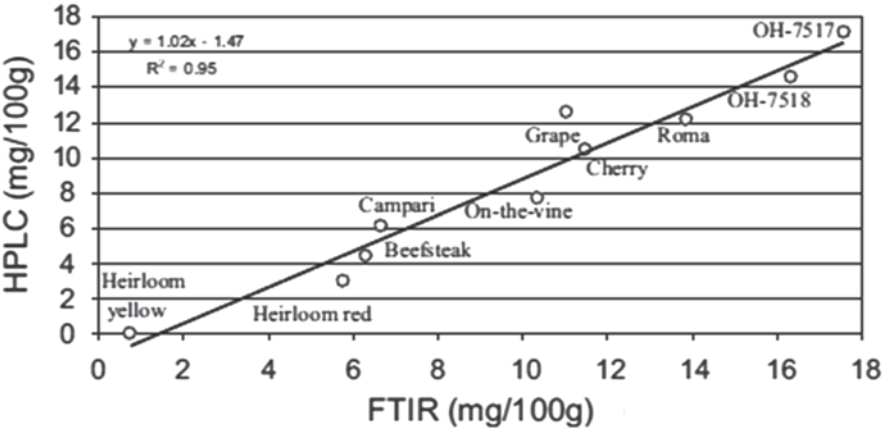


Fig. 10.11 Correlation between lycopene content determined by mid-IR and HPLC in genetically diverse tomatoes. (Adapted with permission from ref [43]. Copyright 2013 American Chemical Society)

Table 10.2 Mid-IR regression statistics of calibration and cross-validation for different carotenoids in apricot flesh. (Reproduced with permission from Ref. [89]. © (ACS Publications) (2013))

Carotenoid compound	λ range (cm^{-1})	LV	Calibration		Cross-validation	
			R^2	RMSEC	R^2	RMSEV
β-cryptoxanthin	940–1,200	11	0.95	0.14	0.22	0.88
γ-carotene	940–1,200	5	0.32	0.26	0.10	0.40
β-carotene	940–1,200	9	0.81	4.08	0.44	9.02
Phytofluene	940–1,200	8	0.68	4.23	0.26	7.63
Phytoene	940–1,200	8	0.75	5.52	0.16	17.42
Total carotenoids	940–1,200	11	0.87	7.62	0.33	25.18
Provitamin A	940–1,200	9	0.90	288.63	0.30	1,533.97

using ATR IR method. It is worth to notice, that the best model was obtained in the limited spectral range (650–1,800 cm^{-1}) because since above 1,800 cm^{-1} IR spectra were disturbed by CO_2 absorption.

However, the properties of biological samples often causes that analysis of natural dyes in many cases is difficult and sometimes even impossible. For example, the calibration and cross-validation of PLS models obtained by ATR IR spectroscopy for carotenoids technique in apricot flesh are shown in the work of Ruiz [89]. After testing of the different wavelength intervals, the spectral range from 940 to 1,200 cm^{-1} was selected for developing the prediction models. Unfortunately, the mid-IR models showed a high validation error (Table 10.2).

The prediction of carotenoids content in apricot turned out not possible by using mid-IR method, probably because of low relative intensity of absorption bands as compared to major compounds such as sugars and acids.

ATR IR combined with a different chemometric method like SIMCA (soft independent modeling by class analogy)—a multivariate analysis technique, which proved to be successful, fast, simple, and robust technique for the profiling of ca-

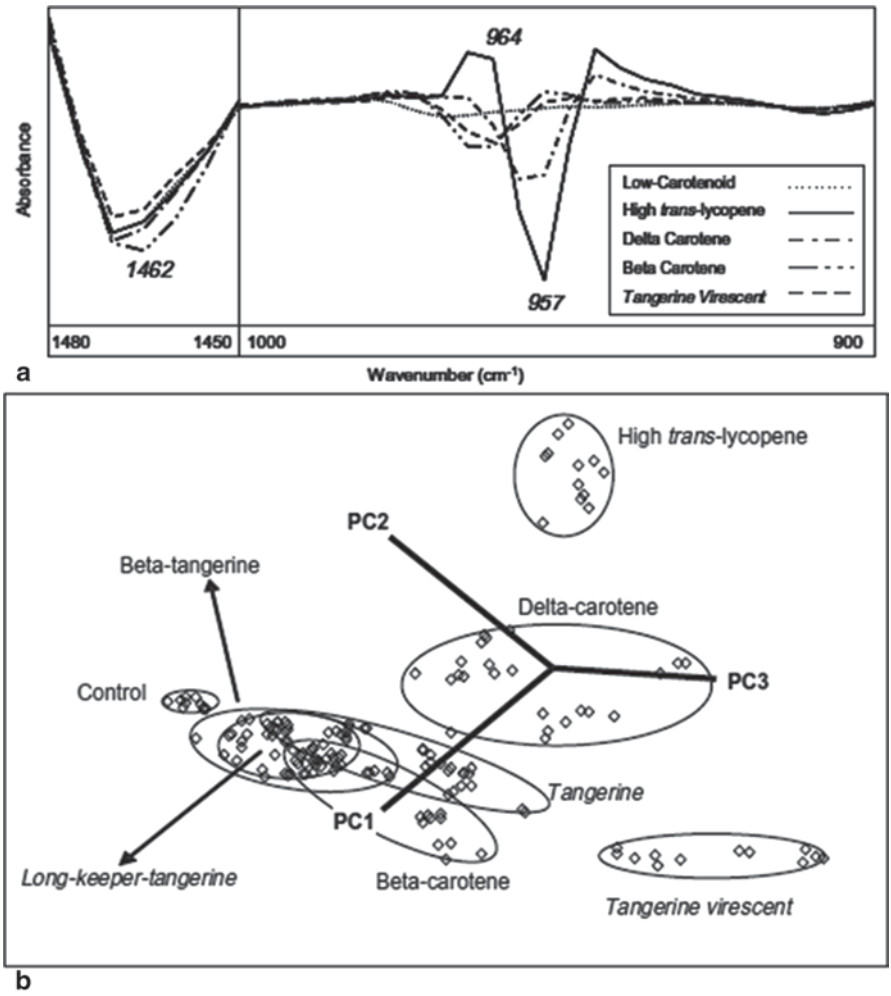


Fig. 10.12 (**a**) Second derivative transformation of ATR IR spectrum of selected tomato varieties in information-rich regions using a ZnSe crystal plate. (**b**) SIMCA tridimensional class projections based on the 850–1,800 cm^{-1} region. (Reproduced with permission from Ref. [90]. © (Elsevier) (2013))

rotenoids in tomatoes [90]. The second derivative of the IR spectra (Fig. 10.12) enabled SIMCA to generate principal component models that differentiated each genotype group from the others based on their predominant carotenoid profile. In this case the identification of differences in the chemical structure of the carotenoids using ATR IR was possible. With the use of this information and SIMCA method one was able to classify each tomato variety by the type and quantity of carotenoids.

The typical example of the application of NIR and chemometry in quantitative analysis of carotenoids has been presented by Pedro and Ferreira [58]. The

studies focused on lycopene and β-carotene content in tomato products with NIR spectrscopy in order to replace standard, time-consuming analytical methods. PLS calibration, based on over one hundred spectra, was chosen. Original spectra were preprocessed by applying a multiplicative signal correction (MSC) or second derivative according to the algorithm proposed by Savitsky and Golay. Prior analysis the spectra were divided into three regions (4,000–6,000 cm^{-1}, 6,000–8,000 cm^{-1} and 8,000–10,000 cm^{-1}). Calibrations were performed in each one, as well as in combinations of them. MSC turned out in the described case the most effective preprocessing technique as well as splitting the spectra into three distinct regions. The results were perfectly comparable with those obtained with HPLC with the "R" values over 0.998.

The mean values obtained by NIR spectroscopy and HPLC for eight carotenoids (α-carotene, β-carotene, α-cryptoxanthin, β-cryptoxanthin, isolutein, lutein, violaxanthin, zeaxanthin) in over sixty maize varieties were equivalent [57, 59]. In order to compare NIR and HPLC results, the modified PLS regression and the lowest standard error of cross-validation were chosen. The correlation factor (R^2) ranged from 0.82 for lutein to 0.94 for zeaxanthin.

Vis/NIR transmittance spectroscopy was successfully used to determine the total levels of carotenoids and chlorophyll also in virgin oil [91]. To correct the signal an initial smoothing technique combined with first derivative treatment was used. Similarly to the former examples PLS regression was used to develop calibration models. Those models were successfully used to monitor on-line levels of the studied compounds during oil processing.

10.3.2 Flavonoids

The *in situ* detection of flavonoids is often disrupted by the overlapping the flavonoid marker bands with signals coming from components of biological matrix e.g. lignin, cellulose.

In 1985, the RR microspectrometry technique was applied to the studies of skins of mature *Vitis vinifera* "Pinot noir" berries and epidermal tissues of the flowers of wild mallow (*Malva sylvestris*) [92]. What is interesting, authors detected malvidin 3-glucoside in "Pinot noir" berries both in the quinonoidal base form (strong band at 1653 cm^{-1}) and in the flavylium form (features at 1,572, 1,600 and 1,648 cm^{-1}).

The Raman mapping has been shown for green rooibos (*Aspalathus linearis*) where the aspalanthin, a dihydrochalcone, is the major flavonoid component [66]. From these results it can be seen that the aspalanthin is distributed heterogeneously in the sample and the highest concentration was observed in the inner part of the leaves. The Raman maps of aspalanthin distribution were obtained according to the intensity of the band at 784 cm^{-1}. The integration over the most intensive bands in the range 1550 cm^{-1}–1,670 cm^{-1} ascribed to the phenyl rings and C=O stretching vibrations and 1,216 cm^{-1} assigned to hydroxyl deformation and ring vibration gives incorrect results because these bands coincide to the signals of the other

components of the plant matrix. Simultaneously, lignin and carotenoids distribution were detected with 1,607 cm^{-1} and 1525 cm^{-1} signals as marker bands, respectively.

The spectacular results of the flavonoid mapping together with carotenoid analyses was published in ref. [71] (Fig. 10.9). The HCA method, calculated for full-range spectra, was used to analyses of the Raman maps that allowed to discriminate within the flavonoids between the flavonols and anthocyanins. Additionally, based on the integrated band intensity, the relative concentrations of these dyes were shown. As the marker bands, two regions were defined: 1,550–1,588 cm^{-1} for flavonols and 1,225–1,272 cm^{-1} for anthocyanins. The sensitivity of HCA allows to discriminate within the flavonoids group between the similar moieties classified to the flavonols and anthocyanins.

Flavonoid components usually do not give so clear and distinctive spectra like e.g. carotenoids. Nevertheless, the *in situ* quantitative analysis is possible. The quantification of aspalathin and nothofagin in green rooibos (*Aspalathus linearis*) was performed with FT-Raman spectroscopy together with flavonoid mapping in plant mentioned above [66]. The FT-Raman calibration for aspalathin model resulted in $R^2 = 0.87$ (SECV = 0.53 g/100 g) whilst for nothofagin showed $R^2 = 0.78$ (SECV—0.14 g/100 g). The last result is relatively satisfactory because nothofagin is presented in small amounts in dried rooibos (0.067–1.23 g/100 g), contrary to aspalathin content which is about 0.60–10.59 g/100 g. In order to apply the PLS regression mode, the whole range of spectrum was selected and the model for the prediction was not improved if only marker bands were calculated.

A few analysis have been carried out for extracted dyes. Quercetin is one of the most popular dietary flavonols occurring in vegetables i.e. onion. The determination of quercetin in the methanol and ethanol extract from an onion peels was carried out by RS with 488 nm laser [93]. What is characteristic, the solvent was used as the internal standard to remove experimental factors disturbing the analysis. For quantitative analysis, the Raman bands of quercetin at 600 cm^{-1} and ethanol at 884 cm^{-1} were used. The analogous analysis was carried for quercetins using ATR FTIR spectroscopy [94].

Detection of anthocyanins and their glycosides [95] was not accompanied by a quantitative analysis and only the spectra of anthocyanins derivatives in dependence of the glycolysation and excitation wavelength was studied.

Lu et al. [96, 97] carried out the FT-IR spectroscopic analysis for onions and shallots, which are characterized by high amount of flavonoids and compared them with the results obtained with the typical analytical methods for phenolic content and total antioxidant capacity (the Folin–Ciocalteu assay, 2,2-diphenyl-picrylhydrazyl (DPPH) assay, Trolox equivalent antioxidant capacity (TEAC) assay and ferric reducing antioxidant power (FRAP) assay). As a result, correlation coefficient > 0.95 was obtained, with standard errors of calibration (SEC) and standard errors of cross-validation (SECV) less than 2.85. Moreover, cluster analysis and discriminant function analysis (DFA) differentiated varieties of onions and shallot based upon infrared spectral features.

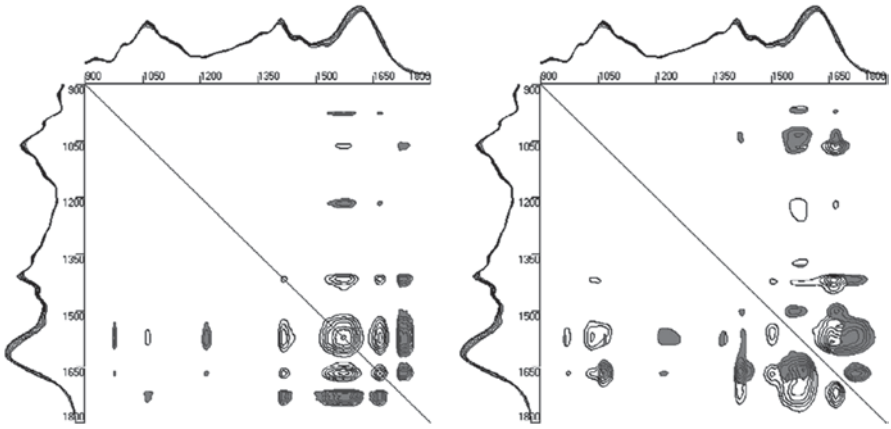


Fig. 10.13 Example of 2D-IR correlation spectra of 'Qing Kai Ling' injection for original sample. *Left*, synchronous spectrum; *right*, asynchronous spectrum (unit, cm^{-1}). (Reproduced with permission from Ref. [98]. © (Elsevier) (2013))

Mid-IR spectroscopy was successfully employed with the multivariate analysis to determine quercetin in onion [94]. The results were compared with HPLC data and the coefficient factors > 0.95 again proved the applicability of both methods.

Two-dimensional correlation infrared transmission spectroscopy allowed the analysis of deterioration of the traditional Chinese medicine, 'Qing Kai Ling', containing flavonoids [98]. According to the authors, in spite of the mid-IR spectra of non-degradated and deteriorated medicine are almost undistinguishable, 2D-IR analysis made this problem possible to solve. Thanks to the perturbation applied to the sample (here, contact with air), the analysis allowed to determine at which positions, in such a measured signal, a systematic changes in spectra were observed. Then, the comparison of 2D correlation spectra of untreated and deteriorated medicine was performed.

In every case, the 2D correlation analysis results in two complementary signals, which referred to the so called 2D synchronous and 2D asynchronous spectra [64]. The both spectra for original Chinese medicine injection are shown at Fig. 10.13.

The analyses showed that the spectra with overlapping band could be unraveled in that case. The most interesting was the wide band at 1,611 cm^{-1} which consisted of three overlapping bands at 1,572, 1,667 and 1,729 cm^{-1}. These bands were assigned among others to the vibrations of the flavone derivatives. Changes in that spectral region brought authors to the conclusion that the main causes of the the deterioration of the studied medicine are the oxidation of flavone compounds and decomposition of glucoside components.

The 2D correlation spectroscopy was applied also by Adiana and Mazura to study the main constituents in the medicinal materials and their extracts from *Senna alata* L.—popular tropical herb [99]. The authors obtained the 2D-IR correlation spectra of extracts by taking advantage of the various response of different chemical

group under thermal perturbation from 50 to 120 °C. Generally, they concluded that the 2D correlation IR spectra not only enlarges the differences in the FT-IR spectra, but also provides useful information for studying the thermal stability of chemical components. The relative contents of flavonoids in the different extracts could be deduced according to the IR characteristic peaks base on the shaping and intensity of peaks near 1,620 cm^{-1}.

Mid-IR turned out also promising techniques in some other dyes analysis. Among others color and antioxidant compounds like phenols, flavons and anthocyanins of different berries were studied [100–102]. The obtained spectral datasets were correlated with technological parameters of ripening or fruit antioxidant activity by using partial least squares (PLS) regression algorithms both in full cross-validation and test set validation. The models developed in the mid-IR region again proved to be well correlated with analytically determinated phenolic compounds.

Similar to the above described, the results for anthocyanins are presented in work of Sinelli et al. [100]. The authors aimed to examine the potential of NIR (as well as mid-IR) spectroscopy to evaluate ripening indices and nutraceutical properties of the two cultivars of blueberries. The spectral datasets obtained were correlated with such parameters like: total soluble solids, total phenols, total flavonoids and total anthocyanins contents by using PLS regression algorithms. In the NIR region the wavenumbers with the highest influence in the PLS models for predicting the parameters above described were 8,859–8,678 cm^{-1}, 7,251–7,154 cm^{-1}, 6,063–5,770 cm^{-1}, 5,345–5,160 cm^{-1}, 4,700–4,500 cm^{-1} assigned to the C-H and O-H overtones and combination bands. As the result NIR methods were characterized by PLS models mainly in full cross-validation by r values from 0.92 to 0.96 for flavonoids and total anthocyanins and turned out very promising alternative quality control method. Moreover, PCA was also successfully performed to evidence separation according to harvest year and a partial separation according to varieties of the blueberries.

NIR method was also applied to determine the total flavonoids content in ginkgo (*Ginkgo biloba* L.) leafs [103]. In traditional Chinese medicine ginkgo leafs were used as a means of treating the symptoms of asthma, cough, and skin infections. Today, extracts from this material, rich in flavonoids, are used in the treatment of diseases of the cardiovascular system as well as supplements for memory and concentration. The total content of flavonoids, however, changes dramatically in fall, which is the period of harvesting. That is why NIR method was used as rapid and nondestructive method of determination of total flavonoids content in ginkgo leaf. Apart from PLS, interval partial least squares (iPLS) and synergy interval partial least squares (SiPLS) were used to develop calibration models.

The authors described the increasing amount of total flavonoids in the following order: green, green-yellowish and yellow leaves. Two characteristic wavelength regions (5,840–6,090 cm^{-1} and 6,620–6,880 cm^{-1}), selected by SiPLS were corresponded to the absorptions of two aromatic ring in basic flavonoid structure (Fig. 10.14). The synergy interval partial least squares method of calibration turned out the most optimal.

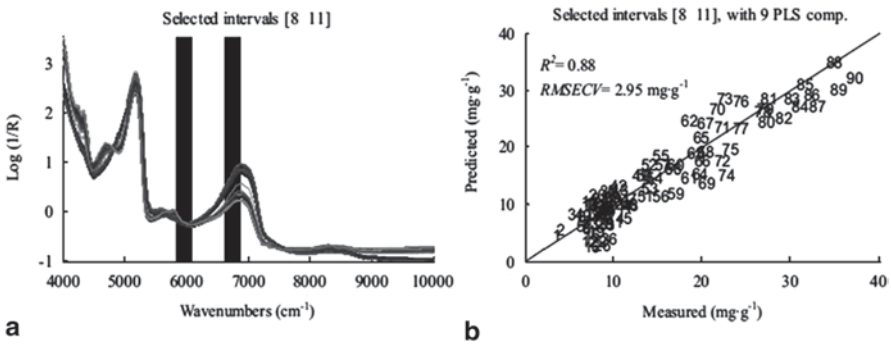


Fig. 10.14 Spectral interval selected by SiPLS (**a**) and reference measured versus NIR prediction by SiPLS in calibration set (**b**) for fresh *Ginkgo biloba* leaves in three colors (green, green-yellow and yellow). (Reproduced with permission from Ref. [103]. © (Elsevier) (2013))

10.3.3 Anthraquinones

In spite of the fact that anthraquinones are widely distributed pigments in nature [6], the vibrational studies of them are not so popular like for the other colorants. The pigments founded in lichens, parietin [85, 104–106] and emodin [104] were detected in natural samples, but quantification analysis was carried out only for parietin accompanied by carotenoids [85]. Generally, the analysis were made with the NIR laser excitation wavelength (1,064 nm).

Most papers about the detection of anthraquinones shows an application of a SERS technique to *in situ* analysis. However, most of them are related to the determination of dye/pigment on the thread in context of the conservation of art [31, 34, 107–109], but not in the biological systems.

10.3.4 Indigo

Indigo shows extremely intense signals in its Raman spectrum that enables to detect this dye directly in textiles and plants [41, 110].

One of the plant species, which contains the indigo precursor, indican, is a dyer's knotweed (*Polygonum tinctorium*). In enzymatically catalyzed process, indigo is formed from indican upon leaf drying and in the case of leaf surface damaging. Both these processes were observed with FT-Raman spectroscopy during two days of drying at room temperature (Fig. 10.15) [41]. The blue area on the leaf surface indicates damages of the plant tissue where the indican was enzymatically transformed to the indigo. The maps were obtained by integration of the marker band near 1,570 cm^{-1}. It is clearly seen that firstly the indigo is present only in the damaged area of a fresh leaf, whilst upon two days, the amount of indigo gradually increased in the leaf area.

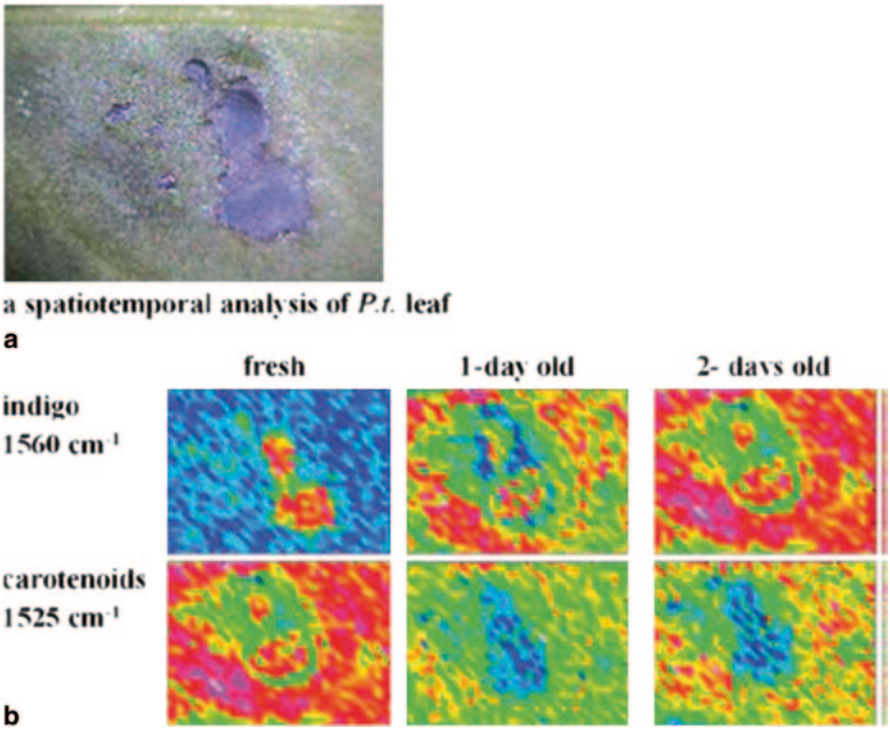


Fig. 10.15 Visual image of *Polygonum tinctorium* leaf (**a**) and FT-Raman maps of indigo and carotenoids distribution (**b**). (*Blue* colour = low concentration, *red* = high concentration of the analyte). (Reproduced with permission from Ref. [41]. © (RSC Publishing) (2013))

Simultaneously, the carotenoids distribution was determined. It is known, that carotenoids act as a plant protector and follow abiotic stress like was demonstrated here, so it was not surprising that those group of dyes was detected.

A very intense Raman signal for indigo pigment enables to analyse it *in situ* without any problems. The same marker band as for mapping, near 1,573 cm^{-1}, was used for developing the quantification protocols for the *in situ* determination of indigo in *Polygonum tinctorium* [111]. The analysis was carried out with FT-Raman spectrometer and HPLC technique as a reference method. The studies of the indigo determination by ATR FTIR spectroscopy did not give any positive results, mainly because of the overlapping of the IR indigo spectrum by the signals of the other plant components e.g. cellulose. In case of FT-RS, indigo signal is so intense that cellulose bands are hardly observed [41, 111].

The first successful synthesis of a synthetic indigo was performed in 1880 by A. von Bayern and since 1897 the commercial production, based on the K. Heumann's synthesis route, were developed [112]. Therefore, especially in the conservation of art textiles, the identification of synthetic or natural indigo enables proper dating of the art objects.

The distinguishing of the synthetic and natural indigo samples is possible with the Raman spectroscopy combined with chemometric methods [112]. The extraction of the essential data was in that case carried out with the PCA as the spectral pre-processing methods whilst the CA and the LDA methods enabled to get the valuable information. This CA result enables to clearly distinguish between natural and synthetic indigo. However, the distinction between natural samples made from *Indigofera tinctoria* and *Isatis tinctoria* was not possible.

The analysis carrying out with the LDA method revealed that the bands at 1,582 cm^{-1} and 1,571 cm^{-1} (ascribed to the symmetric stretching modes of v(C=C), v(C=O and v(C-C)) contribute to the discrimination between the natural and synthetic samples. Moreover, also the region 300–900 cm^{-1} seems to be significant for sample discrimination.

It is also worth to look at the pre-processed spectra prepared for LDA method. To minimize the background influence, the 1st and 2nd derivatives were taken from spectra. The comparison of the results revealed that the 1st derivatives give the worst effect than the untreated and 2nd derivatives spectra.

10.4 Summary

The use of chemometric methods in the analysis of Raman and IR spectra of dyes has opened new possibilities in the areas previously reserved only for chromatographic methods or conventional UV-vis studies. In the future, the combination of chemometrics with spectroscopic methods is likely to replace other measurement techniques for dye analysis, traditionally applied in various fields from food quality to drug control. The rapidity and simplicity together with the possibility of *in situ* measurements are definitely the most attractive feature of these methods.

Acknowledgments AR thanks the project "Society–Environment–Technologies" realized within the Human Capital Operational Programme by Jagiellonian University for financial support.

References

1. De Rijke E, Out P, Niessen WMA, Ariese F et al (2006) Analytical separation and detection methods for flavonoids. J Chromatogr A 1112:31–63
2. Melo MJ (2009) History of natural dyes in the ancient mediterranean world. In: Bechtold T, Mussak R (eds) Handbook of natural colorants. Wiley, Chichester
3. Stuart BH (2007) Molecular spectroscopy. Analytical techniques in materials conservation. Wiley, Chichester
4. Zollinger H (2003) Color chemistry: syntheses, properties, and applications of organic dyes and pigments. Wiley, Chichester
5. General Introduction to the Chemistry of Dyes (2008) Some aromatic amines, organic dyes, and related exposures. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans

6. Delgado-Vargas F, Jiménez AR, Paredes-López O (2000) Natural pigments: carotenoids, anthocyanins, and betalains-characteristics, biosynthesis, processing, and stability. Crit Rev Food Sci Nutr 40:173–289
7. Schmidt W (2005) Optical spectroscopy in chemistry and life science. Wiley-VCH, Weinheim
8. Wissgott U, Bortlik K (1996) Prospects for new natural food colorants. Trends Food Sci Technol 7:298–302
9. Schulz H, Baranska M (2007) Identification and quantification of valuable plant substances by IR and Raman spectroscopy. Vib Spec 43:13–25
10. Baranska M, Schütz W, Schulz H (2006) Determination of lycopene and b-carotene content in tomato fruits and related products: comparison of FT-Raman, ATR-IR, and NIR spectroscopy. Anal Chem 78:8456–8461
11. Schulz H, Baranska M, Baranski R (2005) Potential of NIR-FT-Raman spectroscopy in natural carotenoid analysis. Biopolymers 77:212–221
12. Siefermann-Harms D (1987) The light-harvesting and protective functions of carotenoids in photosynthetic membranes. Physiol Plant 69:561–568
13. Maia LF, de Oliveira VE, de Oliveira MER et al (2012) Polyenic pigments from the Brazilian octocoral Phyllogorgia dilatata Esper, 1806 characterized by Raman spectroscopy. J Raman Spectrosc 43:161–164
14. Pini E, Bertelli A, Stradi R, Falchi M (2004) Biological activity of parrodienes, a new class of polyunsaturated linear aldehydes similar to carotenoids. Drugs Exp Clin Res 30:203–206
15. McGraw KJ, Nogare MC (2004) Carotenoid pigments and the selectivity of psittacofulvin-based coloration systems in parrots. Comp Biochem Physiol B Biochem Mol Biol 138:229–233
16. Bechtold T (2009) Natural colorants—quinoid, naphthoquinoid and anthraquinoid dyes. In: Bechtold T, Mussak R (eds) Handbook of natural colorants. Wiley, Chichester
17. Giusti MM, Wallace TC (2009) Flavonoids as natural pigments. In: Bechtold T, Mussak R (eds) Handbook of natural colorants. Wiley, Chichester
18. Harborne JB, Willimas CA (2000) Advances in flavonoid research since 1992. Phytochem 55:481–804
19. Lakhanpal P, Rai DK (2007) Quercetin: a versatile flavonoid. Internet J Med Update 2:22–37
20. Schlangen K, Miosic S, Castro A et al (2009) Formation of UV-honey guides in Rudbeckia hirta. Phytochem 70:889–898
21. Andersen ØM, Markham KR (2006) Flavonoids chemistry, biochemistry and applications. CRC Press, New York
22. Cockell CS, Knowland J (1999) Ultraviolet radiation screening compounds. Biol Rev 74:311–345
23. Markham KR, Gould KS, Ryan KG (2001) Cytoplasmic accumulation of flavonoids in flower petals and its relevance to yellow flower colouration. Phytochemistry 58:403–413
24. John P, Angelini LG (2009) Indigo—agricultural aspects. In: Bechtold T, Mussak R (eds) Handbook of natural colorants. Wiley, Chichester
25. Schweppe H (1993) Handbuch der Naturfarbstoffe. Nikol Verlagsgesellschaft, Hamburg
26. Cooksey CJ (2001) Tyrian Purple: 6,6'-Dibromoindigo and related compounds. Molecules 6:736–769
27. Schrader B, Klump HH, Schenzel K, Schulz H (1999) Non-destructive NIR-FT Raman analysis of plants. J Mol Stuct 509:201–212
28. Schrader B, Schulz H, Andreev GN et al (2000) Non-destructive NIR-FT-Raman spectroscopy of plant and animal tissues, of food and works of art. Talanta 53:35–45
29. Dietzek B, Cialla D, Schmitt M, Popp J (2010) Introduction to the fundamentals of Raman spectroscopy. In: Dieing T, Hollricher O, Toporski J (eds) Confocal Raman spectroscopy. Springer, Berlin
30. Aroca R (2006) Surface-enhanced vibrational spectroscopy. Wiley, Chichester
31. Jurasekova Z, del Puerto E, Bruno G et al (2010) Extractionless non-hydrolysis surface-enhanced Raman spectroscopic detection of historical mordant dyes on textile fibers. J Raman Spectrosc 41:1455–1461

32. Bruni S, Gugliemi V, Pozzi F (2011) Historical organic dyes: a surface-enhanced Raman scattering (SERS) spectral database on Ag Lee–Meisel colloids aggregated by NaClO4. J Raman Spectrosc 42:1267–1281
33. Pozzi F, Lombardi JR, Bruni S, Leona M (2012) Sample treatment considerations in the analysis of organic colorants by surface-enhanced Raman scattering. Anal Chem 84:3751–3757
34. Casadio F, Leona M, Lombardi JR, Van Duyne R (2010) Identification of organic colorants in fibers, paints, and glazes by surface enhanced Raman spectroscopy. Acc Chem Res 43:782–791
35. Hering K, Cialla D, Ackermann K et al (2008) SERS: a versatile tool in chemical and biochemical diagnostics. Anal Bioanal Chem 390:113–124
36. Madariaga JM (2012) Identification of dyes and pigments by vibrational spectroscopy. In: Chalmers JM, Edwards HGM, Hargreaves MD (eds) Infrared and Raman Spectroscopy in forensic science. Wiley, Chichester
37. Dent G (2006) Vibrational spectroscopy of colors dyes and pigments In: Griffiths P, Chalmers JM (eds) Handbook of vibrational spectroscopy. Wiley, Chichester
38. Mirabella FM (2006) Principles, theory and practice of internal reflection spectroscopy. In: Griffiths P, Chalmers JM (eds) Handbook of vibrational spectroscopy. Wiley, Chichester
39. Kokot S, Crawford K, Rintoul L, Meyer U (1997) A DRIFTS study of reactive dye states on cotton fabric. Vib Spectrosc 15:103–111
40. Navas N, Romero-Pastor J, Manzano E, Cardell C (2008) Benefits of applying combined diffuse reflectance FTIR spectroscopy and principal component analysis for the study of blue tempera historical painting. Anal Chim Acta 630:141–149
41. Baran A, Fiedler A, Schulz H, Baranska M (2010) In situ Raman and IR spectroscopic analysis of indigo dye. Anal Methods 2:1372–1376
42. Bruni S, Caglio S, Guglielmi V, Poldi G (2008) The joined use of n.i. spectroscopic analyses—FTIR, Raman, visible reflectance spectrometry and EDXRF—to study drawings and illuminated manuscripts. Appl Phys A 92:103–108
43. Rubio-Diaz DE, Francis DM, Rodriguez-Saona LE (2011) External calibration models for the measurement of tomato carotenoids by infrared spectroscopy. J Food Comp Anal 24:121–126
44. Casadio F, Toniolo L (2001) The analysis of polychrome works of art: 40 years of infrared spectroscopic investigations, J Cult Herit 2:71–78
45. Wang XF, Yu J, Zhang AL et al (2012) Nondestructive identification for red ink entries of seals by Raman and Fourier transform infrared spectrometry. Spectrochim Acta A 97:986–994
46. Koperska M, Lojewski T, Lojewska J (2011) Vibrational spectroscopy techniques to study degradation of natural dyes. Assessment of oxygen-free cassette for safe exposition of artefacts. Anal Bioanal Chem 399:3271–3283
47. Miliani C, Domenici D, Clementi C et al (2012) Colouring materials of pre-Columbian codices: non-invasive in situ spectroscopic analysis of the Codex Cospi. J Archaeol Sci 39:672–679
48. Saurina J (2010) Characterization of wines using compositional profiles and chemometrics. TrAC-Trend Anal Chem 29:234–245
49. Pappas CS, Takidelli C, Tsantili E et al (2011) Quantitative determination of anthocyanins in three sweet cherry varieties using diffuse reflectance infrared Fourier transform spectroscopy. J Food Comp Anal 24:17–21
50. Bauer R, Nieuwoudt H, Bauer FF et al (2008) FTIR spectroscopy for grape and wine analysis. Anal Chem 80:1371–1379
51. Strachan CJ, Rades T, Gordon KC, Ranten J (2007) Raman spectroscopy for quantitative analysis of pharmaceutical solids. J Pharm Pharmacol 59:179–192
52. Blanco M, Villarroya I (2002) NIR spectroscopy: a rapid-response analytical tool. TrAC-Trends Anal Chem 21:240–250
53. Davey MW, Saeys W, Hof E et al (2009) Application of Visible and Near-Infrared Reflectance Spectroscopy (Vis/NIRS) to determine carotenoid contents in banana (Musa spp.) fruit pulp. J Agric Food Chem 57:1742–1751
54. Sharpe PJH, Barber HN (1972) Near infrared reflectance of colored fruits. Appl Opt 11:2902–2906
55. Næs T, Isaksson T, Fearn T, Davies T (2002) A user-friendly guide to multivariate calibration and classification. NIR Publications, Chichester

56. Reich G (2005) Near-infrared spectroscopy and imaging: basic principles and pharmaceutical applications. Adv Drug Deliv Rev 57:1109–1143
57. Berardo N, Brenna OV, Amato A, Valotia P, Pisacanea V, Mottoa M (2004) Carotenoids concentration among maize genotypes measured by near infrared reflectance spectroscopy (NIRS). Innov Food Sci Emerg 5:393–398
58. Pedro AM, Ferreira MM (2005) Nondestructive determination of solids and carotenoids in tomato products by near-infrared spectroscopy and multivariate calibration. Anal Chem 77:2505–2511.
59. Brenna OV, Berardo N (2004) Application of Near-Infrared Reflectance Spectroscopy (NIRS) to the evaluation of carotenoids content in maize. J Agric Food Chem 52:5577–5582.
60. Treado PJ, Nelson MP (2002) Raman imaging. In: Chalmers JM, Griffiths PR (eds) Handbook of vibrational spectroscopy. John Wiley & Sons Ltd., Chichester
61. Gierlinger N, Schwanninger M (2007) The potential of Raman microscopy and Raman imaging in plant research—review. Spectroscopy 21: 69–89
62. Dieing T, Ibach W (2010) Raman microscopy. In: Dieing T, Hollricher O Toporski J (eds) Confocal Raman microscopy. Springer-Verlag Berlin Heidelberg
63. De Juan A, Maeder M, Hancewicz T, Duponchel L, Tauler R (2009) Chemometric tools for image analysis. In: Salzer R, Siesler HW (eds) Infrared an Raman spectroscopic imaging. Wiley-VCH, Chichester
64. Noda I, Ozaki Y (2004) Two-Dimensional correlation spectroscopy: Applications in vibrational and optical spectroscopy. Wiley, Chichester
65. Brereton RG (2007) Applied chemometrics for scientists. Wiley, Chichester
66. Baranska M, Schulz H, Joubert E, Manley M (2006) In situ flavonoid analysis by FT-Raman spectroscopy: identification, distribution, and quantification of aspalathin in green rooibos (Aspalathus linearis). Anal Chem 78:7716–7721
67. Baranski R, Baranska M, Schulz H (2005) Changes in carotenoid content and distribution in living plant tissue can be observed and mapped in situ using NIR-FT-Raman spectroscopy. Planta 222:448–457
68. Withnall R, Chowdry BZ, Silver J, Edwards HGM, de Oliveira LFC (2003) Raman spectra of carotenoids in natural products. Spectrochim Acta A 59:2207–2212
69. Veronelli M Zerbi G, Stradi R (1995) In situ resonance Raman spectra of carotenoids in bird's feathers. J Raman Spectrosc 26:683–692
70. De Oliveira VE, Castro HV, Edwards HGM, de Oliveira LFC (2010) Carotenes and carotenoids in natural biological samples: a Raman spectroscopic analysis. J Raman Spectrosc 41:642–650
71. Gamsjaeger S, Baranska M, Schulz H, Heiselmayer P, Musso M (2011) Discrimination of carotenoid and flavonoid content in petals of pansy cultivars (Viola x wittrockiana) by FT-Raman spectroscopy. J Raman Spectrosc 42:1240–1247
72. Kaczor A, Turnau K, Baranska M (2011) In situ Raman imaging of astaxanthin in a single microalgal cell. Analyst 136:1109–1112
73. Huang YY, Beal CM, Cai WW, Ruoff RS, Terentjev EM (2010) Micro-Raman spectroscopy of algae: composition analysis and fluorescence background behavior. Biotechnol Bioeng 105:889–898
74. Bartlett H, Howells O, Eperjesi F (2010) The role of macular pigment assessment in clinical practice: a review. Clin Exp Optom 93:300–308
75. Bernstein PS, Zhao D-Y, Sharifzadeh M, Ermakov IV, Gellermann W (2004) Resonance Raman measurement of macular carotenoids in the living human eye. Arch Biochem Biophys 430:163–169
76. Tanito M, Obana A, Gohto Y, Okazaki S, Gellermann W, Ohira A (2012) Macular pigment density changes in Japanese individuals supplemented with lutein or zeaxanthin: quantification via resonance Raman spectrophotometry and autofluorescence imaging. Jpn J Ophthalmol 56:488–496
77. Howells O, Eperjesi F, Bartlett H (2011) Measuring macular pigment optical density in vivo: a review of techniques. Graefes Arch Clin Exp Ophthalmol 249:315–347

78. Casella M, Lucotti A, Tommasini M, Bedoni M, Forvi E, Gramatica F, Zerbi G (2011) Raman and SERS recognition of β-carotene and haemoglobin fingerprints in human whole blood. Spectrochim Acta Part A 79:915–919
79. Darvin ME, Gersonde I, Meinke M, Sterry W, Lademann J (2005) Non-invasive in vivo determination of the carotenoids beta-carotene and lycopene concentrations in the human skin using the Raman spectroscopic method. J Phys D: Appl Phys 38:2696–2700
80. Maia LF, Fleury BG, Lages BG, Barbosa JP, Pinto AC, Castro HV, de Oliveira VE, Edwards HGM, de Oliveira LFC (2011) Identification of reddish pigments in octocorals by Raman spectroscopy. J Raman Spectrosc 42:653–658
81. Hedegaard C, Bardeau J-F, Chateigner D (2006) Molluscan shell pigments: an in situ resonance Raman study. J Molluscan Stud 72:157–162
82. Barnard W, de Waal D (2006) Raman investigation of pigmentary molecules in the molluscan biogenic matrix. J Raman Spectrosc 37:342–352
83. Anastasaki EG, Kanakis HD, Pappas C, Maggi L, Zalacain A, Carmona M, Alonso GL, Polissiou MG (2010) Quantification of crocetin esters in saffron (Crocus sativus L.) using Raman spectroscopy and chemometrics. J Agric Food Chem 58:6011–6017
84. Bhosale P, Ermakov IV, Ermakova MR, Gellermann W, Bernstein PS (2004) Resonance Raman quantification of nutritionally important carotenoids in fruits, vegetables, and their juices in comparison to high-pressure liquid chromatography analysis. J Agric Food Chem 52:3281–3285
85. Edwards HGM, Cockell CS, Newton EM, Wynn-Williams DD (2004) Protective pigmentation in UVB-screened Antarctic lichens studied by Fourier transform Raman spectroscopy: an extremophile bioresponse to radiation stress. J Raman Spectrosc 35:463–469
86. Doka O, Bicanic D, Buijnsters JG, Spruijt R, Luterotti S Vegvari G (2010) Exploiting direct and indirect methods for the detection of the total carotenoid content in dried pastas. Eur Food Res Technol 230:813–819
87. Lunde K, Zechmeister L (1955) Infrared Spectra and cis-trans configurations of some carotenoid pigments. J Am Chem Soc 77:1647–1653
88. De Nardo T, Shiroma-Kian C, Halim Y, Francis D, Rodriguez-Saona L (2009) Rapid and simultaneous determination of lycopene and beta-carotene contents in tomato juice by infrared spectroscopy. J Agr Food Chem 57:1105–1112
89. Ruiz D, Reich M, Bureau S, Renard CM, Audergon JM (2008) Application of reflectance colorimeter measurements and infrared spectroscopy methods to rapid and nondestructive evaluation of carotenoids content in apricot (Prunus armeniaca L.). J Agr Food Chem 56:4916–4922
90. Rubio-Diaz DE, De Nardo T, Santos A, de Jesus S, Francis D Rodriguez-Saon LE (2010) Profiling of nutritionally important carotenoids from genetically-diverse tomatoes by infrared spectroscopy. Food Chem 20:282–289
91. Marquez AJ (2003) Monitoring carotenoid and chlorophyll pigments invirginolive oil by visible-near infrared transmittance spectroscopy. On-line application. J Near Infrared Spec 11:219–226
92. Merlin J-C, Statoua A Brouillard R (1985) Investigation of the in vivo organization of anthocyanins using resonance Raman microspectrometry. Phytochem 24:1575–1581
93. Numata Y, Tanaka H (2011) Quantitative analysis of quercetin using Raman spectroscopy. Food Chemistry 126:751–755
94. Lu X, Ross CF, Powers JR, Rasco BA (2011) Determination of quercetins in onion (Allium cepa) using infrared spectroscopy. J Agric Food Chem 59:6376–6382
95. Merlin JC, Statoua A, Cornard JP, Saidi-Idrissi M, Brouillard R (1994) Resonance Raman spectroscopic studies of anthocyanins and anthocyanidins in aqueous solutions. Phytochem 35:227–232
96. Lu X, Wang J, Hamzah M Al-Qadiri HM, Ross CF, Powers JR, Tang J, Rasco BA (2011) Determination of total phenolic content and antioxidant capacity of onion (Allium cepa) and shallot (Allium oschaninii) using infrared spectroscopy Food Chem 129:637–644

97. Lu X, Ross CF, Powers JR, Aston DE, Rasco BA (2011) Determination of total phenolic content and antioxidant activity of garlic (Allium sativum) and elephant garlic (Allium ampeloprasum) by attenuated total reflectance—Fourier transformed infrared spectroscopy. J Agr Food Chem 59:5215–5221
98. Zuo L, Sun Sq, Zhou Q, Tao Jx, Noda I (2003) 2D-IR correlation analysis of deteriorative process of traditional Chinese medicine 'Qing Kai Ling' injection. J Pharm Biomed Anal 30:1491–1498
99. Adiana MA Mazura MP (2011) Study on Senna alata and its different extracts by Fourier transform infrared spectroscopy and two-dimensional correlation infrared spectroscopy. J Mol Struct 991:84–91
100. Sinelli N, Spinardi A, Di Egidio V, Mignani I Casiraghi E (2008) Evaluation of quality and nutraceutical content in blueberries (Vaccinium corymbosum L.) by near and mid-infrared spectroscopy. Postharvest Biol Tec 50:31–36
101. Kemsley EK, Holland JK, Defernez M Wilson RH (1996) Detection of adulteration of raspberry purees using infrared spectroscopy and chemometrics. J Agric Food Chem 44:3864–3870
102. Lam HS, Proctor A, Howard L Cho MJ (2005) Rapid fruit extracts antioxidant capacity determination by Fourier transform infrared spectroscopy. J Food Sci 70:C545–C549
103. Shi JY, Zou XB, Zhao JW, Mel H, Wang KL, Wang X Chen H (2012) Determination of total flavonoids content in fresh Ginkgo biloba leaf with different colors using near infrared spectroscopy. Spectrochim Acta A 94:271–276
104. Edwards HGM, Newton EM, Wynn-Williams DD, Coombes SR (2003) Molecular spectroscopic studies of lichen substances. 1. Parietin and emodin. J Mol Struct 648:49–59
105. Edwards HGM, Newton EM, Wynn-Williams DD, Dickensheets D, Schoen C, Crowder C (2003) Laser wavelength selection for Raman spectroscopy of microbial pigments in situ in Antarctic desert ecosystem analogues of former habitats on Mars. Int J Astrobiology 1:333–348
106. Edwards HGM, Newton EM, Wynn-Williams DD, Lewis-Smith RI (2003) Non-destructive analysis of pigments and other organic compounds in lichens using Fourier-transform Raman spectroscopy: a study of Antarctic epilithic lichens. Spectrochim Acta A 59:2301–2309
107. Bruni S, Gugliemi V, Pozzi F (2010) Surface-enhanced Raman spectroscopy (SERS) on silver colloids for the identification of ancient textile dyes: Tyrian purple and madder. J Raman Spectrosc 41:175–180
108. Lofrumento C, Ricci M, Platania E, Beucci M, Castellucci E (2013) SERS detection of red organic dyes in Ag-agar gel. J Raman Spectrosc 44:47–54
109. Leona M, Decuzzi P, Kubic TA, Gates G, Lombardi JR (2011) Nondestructive identification of natural and synthetic organic colorants in works of art by surface enhanced Raman scattering. Anal Chem 83:3990–3993
110. Coupry C, Sagon G, Goruet-Ballesteros P (1997) Raman spectroscopic investigation of blue contemporary textiles. J Raman Spectrosc 28:85–89
111. Fiedler A, Baranska M, Schulz H (2011) FT-Raman spectroscopy—a rapid and reliable quantification protocol for the determination of natural indigo dye in Polygonum tinctorium. J Raman Spectrosc 42:551–557
112. Vandenabeele P, Moens L (2003) Micro-Raman spectroscopy of natural and synthetic indigo samples. Analyst 128:187–193

Chapter 11
Structural and Spatial Analysis of Carotenoids in a Single Cell Monitored by Raman Spectroscopy

Agnieszka Kaczor and Marta Pilarczyk

Abstract Carotenoids are particularly convenient subjects to study by means of Raman spectroscopy due to very high Raman scattering cross-section associated with their chain vibrations. Raman studies of carotenoids in single cells have a variety of applications starting from monitoring of growth and morphogenesis of unicellular algae and ending with differentiation of cancerous versus non-cancerous tissues in humans. Examples illustrating the potential of Raman spectroscopy to investigate carotenoids structure and distribution with a particular impact on studies using chemometric and computational methods as a tool to analyze experimental data are reviewed in this chapter.

Keywords Carotenoids • Raman • Imaging • Chemometrics • Quantum-chemical calculations • Unicellular algae

11.1 Introduction

Raman spectroscopy is a particularly convenient method to study carotenoids due to the long chain of alternating double and single bonds in their structures, responsible for unusually high Raman scattering cross-section associated with the chain vibrations. Additionally, upon excitation in the visible absorption range, the resonance effect is observed for carotenoids resulting in an enhancement factor of about five orders of magnitude compared to nonresonant conditions [1]. It makes Raman spectroscopy suitable technique to study carotenoids in low-concentration conditions such as human skin [2], eye [3], human cells [4–6] or immune system [7–10]. Such biochemical studies gain significantly if the pigment localization in a cell or

A. Kaczor (✉) · M. Pilarczyk
Faculty of Chemistry, Jagiellonian University, Ingardena 3, 30-060 Krakow, Poland
e-mail: kaczor@chemia.uj.edu.pl

A. Kaczor
Jagiellonian Centre of Experimental Therapeutics (JCET), Jagiellonian University, Bobrzynskiego 14, 30-348 Krakow, Poland

M. Baranska (ed.), *Optical Spectroscopy and Computational Methods in Biology and Medicine,* Challenges and Advances in Computational Chemistry and Physics 14,
DOI 10.1007/978-94-007-7832-0_11, © Springer Science+Business Media Dordrecht 2014

tissue can be accurately determined. In this context, a Raman spatially sensitive technique—Raman imaging emerges a particularly promising tool of study.

Raman imaging is a non-invasive technique of studying components distribution and concentration in a sample. The technique combines the structural specificity of Raman spectroscopy with the high spatial resolution of microscopy (lateral resolution of 360 nm and depth resolution of *ca*. 1 μm @ 532 nm excitation) resulting in possibility of measuring sample microstructure at the subcellular level. 3D images of tissue and cellular components, illustrating high heterogeneity of the studied objects, have been already obtained with the application of confocal Raman microimaging [11].

Due to the complexity and heterogenity of studied biochemical samples, data analysis is usually supported by chemometrics. Chemometrics is mathematics and statistics based discipline allowing for interpretation of complicated datasets, such as the Raman spectral representation of a sample. Classical, particularly regression-based chemometric methods concentrated on exact identification of the sample composition. This approach, although valid for simple systems, does not work particularly well for biological samples of considerable heterogeneity and complex spatial distribution. In such cases, data analysis aims rather at determination of chemical markers that describe changes in the system [12]. Depending on the studied system and research aim, both chemometric approaches were applied in carotenoid analysis of biological systems. For example, regression-based methods were used to classify carotenoid-producing unicellular organisms of different type or in modified environment (nitrogen-replete and nitrogen-starved conditions) [13–15]. In the "marker-aimed" approach, the application of K-means algorithm to analysis of Raman images of cancerous versus non-cancerous tissues allowed for determination of molecular markers of the disease with a reduced carotenoid level being one of them [6]. This potential of Raman spectroscopy to distinguish between cancerous and non-cancerous tissues/cells is a hope for development of non-invasive Raman-based systems for medical diagnostics (including biopsy-free diagnosis), although routine clinical usage of Raman spectroscopy is still a song of the future.

Another potentially promising field of application of Raman spectroscopy is monitoring of growth and morphogenesis of unicellular algae. Unicellular algae, some species of yeast, bacteria and fungi are natural *de novo* producers of carotenoids and, therefore, a potent biosource of these chemicals [16]. Additionally, some algal strains are potential biofuel source due to massive lipid production (oil content in the range of 20–50 % dry weight of biomass) [17]. Therefore, it is not surprising that already in 2004, the global market for microalgal biomass was estimated to be 5,000 t of dry matter per year [18]. Raman spectroscopy is an effective tool to optimize conditions of carotenogenesis and lipid production by algal cells in laboratory conditions and, undoubtedly, has a potential to became a routine control system in industrial, large-scale algal cultures. As carotenogenesis processes in pigment producing unicellular organisms are not fully understood, there is a large field of application of Raman spectroscopy to study these processes both qualitatively and quantitatively, while Raman imaging can be used to spatially localize produced pigments. Chemometric methods are often used to recognize between synthesized

carotenoids of similar vibrational signature [19], distinguish cells grown in different environmental conditions [13–15], and find differences between carotenoid composition in different species [15]. Possibility of the application of Raman microspectroscopy to study a single cell rather than bulk sample is of high importance due to observed heterogeneity of cellular populations in unicellular species [20, 21].

Quantum-chemical calculations have been formerly applied to solve various structure-related problems related to carotenoids, such as reaction dynamics of carotenoids as antioxidants [22], various aspects of carotenoid reactivity in light harvesting complexes [23–26], or a batochromic shift of protein-bound carotenoid astaxanthin [27, 28]. It is, however, necessary to underline that carotenoids, 40 carbon atoms molecules, are computationally demanding even with the application of powerful computing resources and fast DFT algorithms. The additional complication, difficult to account for in computations, is a significant change of pigment structures in biological environment as demonstrated for instance by abovementioned significant batochromic shift of protein-bound astaxanthin compared to a non-bonded standard [27, 28]. Therefore, the quantum-chemical modeling of carotenoids in the biological context, including cells and tissues, is rather scarce. Examples of such studies, illustrating the potential of quantum-chemical methods to successfully solve structural problems of bound-carotenoids, are investigations of the structure and distribution of astaxanthin in a unicellular alga *Haematococcus Pluvialis*, also upon the temperature stress [29, 30].

The chapter is divided in three parts. The first section is devoted to studies of structure and cellular distribution of carotenoids in unicellular organisms, producing these pigments *de novo*. The second section covers works related to the application of Raman spectroscopy to study carotenoids in plant and animal (non-human) cells and finally, the third section reviews research in which Raman spectroscopy and imaging were used to analyze these pigments in human healthy and pathogenic cells.

11.2 Raman Imaging of Unicellular Organisms

Considerable attention has been focused on carotenoids in unicellular organisms producing these pigments *de novo* with Raman microspectroscopy and microimaging being very effective methods to study these micro-scale systems. Some of the works in this area are purely experimental and they will be referred briefly below. Examples of the successful match between experimental and computational (both in the sense of the application of quantum-chemical and chemometric methods) strategies in the studies of unicellular organisms are described at the end of the chapter.

Kubo et al. [31] applied Raman mapping and polarized Raman spectroscopy to determine distribution of pigment molecules and their orientation, respectively, in a single cell of *Euglena gracilis* and *Chlamydomonas reinhardtii*. The observed Raman signal was derived solely from carotenoids, although the pigment(s) were not specified. Integration of the marker band at 1530 cm^{-1} allowed for localization of

the areas of high pigment intensity in the *Euglena* and *Chlamydomonas* cells. Both for *Euglena* and *Chlamydomonas*, the highest abundance of carotenoids were found in the eyespot, however, a pigment was also observed in other areas of the cell, most probably in chloroplasts. In *Chlamydomonas*, due to known birefringence of its eye, the eyespot was measured with the polarized light, showing that pigment molecules are arranged with the long axis along the elongated direction of the eyespot and parallelly to the body axis [31].

Raman and fluorescence microspectroscopy, in addition to matrix-free laser desorption/ionization mass spectrometry were applied to the analysis of composition and distribution of photosynthetic apparatus components in a single cell of *Euglena gracilis* and *Chlamydomonas reinhardtii* [32]. High-spatial resolution (*ca.* 550 nm) Raman imaging combined with fluorescence imaging allowed for simultaneous visualization of β-carotene and protoplastids distribution in a single algal cell.

In the light of increasing interest in the application of algae as biofuel, two algal species, a potential biofuel source, *Chlorella sorokiniana* and *Neochloris oleoabundans*, were investigated with the help of Raman microimaging [33]. The compositional analysis of healthy and nitrogen-starved algae indicated for the presence of an unspecified carotenoid in their cells, accompanied with chlorophyll and triglyceride in the case of nitrogen-starved algae. Raman distribution maps showed that the unspecified carotenoid is located specifically within a chloroplast [33].

Laser tweezers Raman spectroscopy (LTRS) was applied to study the carotenogenesis process in a single live cell of yeast *Rhodotorula glutinis* [20]. The LTRS-based analysis of the kinetics of total carotenoid accumulation inside individual *Rhodotorula glutinis* cells demonstrated a significant variation of the carotenoid level in the studied population of 100 cells, particularly for certain growth phases. This variation indicates for high cellular heterogeneity in the studied cellular population, and provide for better understanding of carotenogenesis not only in the bulk sample, but for individual cells [20].

Raman microspectroscopy was also demonstrated to be a suitable tool to quantify carotenoids inside subcellular structures. β-carotene quantification in lipid bodies of the alga *Trachydiscus minutus in vivo* was assessed and proposed to be a method for noninvasive cell sorting of algal strains [21]. Calculated β-carotene absolute concentrations were correlated with lipid bodies volumes, showing high variability between the individual specimens.

Spontaneous Raman spectroscopy was also combined with Coherent anti-Stokes Raman Spectroscopy (CARS) to study carotenoid-related lipid accumulation under nitrogen-repletion conditions in a green microalga *Coccomyxa subellipsoidea* [34]. As mentioned above, such research are of growing interest due to potential industrial usage of Raman spectroscopy as a tool for quantification of biofuels or chemicals in algal cells. As expected, both spontaneous Raman and CARS imaging demonstrated that the lipid (triacylglycerols) level increased in N-depleted algae compared with control algal cells containing more carotenoids and less lipids. In the case of the applied equipment, CARS spectroscopy was demonstrated to possess higher scanning speed and spatial resolution (<500 nm) [34], although it is necessary to

mention that such resolution is theoretically achievable for the conventional Raman imaging with the applied excitation wavelength (514 nm).

Recently, a new approach eliminating CARS complications related to non-resonant background and phase matching conditions, i. e. Multiplex Stimulated Raman Scattering Microscopy (SRS) has been applied to study distribution and concentration of main components of the *Botryococcus Braunii* microalga grown upon different lighting conditions [35]. The application of SRS resulted in obtaining, in a fast manner, spatially-resolved images of chlorophyll, carotenoids, lipids and proteins and showed that pigments accumulate in the periphery of the cell. A significant increase of carotenoids level (by 95 %) in the conditions of continuous illumination compared to 12:12 light:dark cycle was noticed, probably due to the increase of the photosynthetic activity of the algal cells upon full light conditions [35].

Works described above were based purely on the experimental approach. Below the example of successful match between experiment and quantum-chemical calculations will be presented in the analysis of astaxanthin in a single algal cell. Astaxanthin (3,3′-dihydroxy-β,β-carotene-4,4′-dione) is announced a superpotent antioxidant, and it is probably the most frequently studied carotenoid in natural sources such as algal cells. The most potent natural bioreactor for astaxanthin production is *Haematococcus Pluvialis*, a unicellular algae that upon starving conditions may accumulate up to 3–4 % of its total dry weight [36]. To obtain information about astaxanthin distribution and structure in a single *Haematococcus* cell, Raman microimaging was applied to study this pigment in the cyst algal cells (aplanospores) [29]. The Raman distribution maps demonstrate that with the maturation of cysts, astaxanthin distribution changes with the pigment shifting from the very center of cells toward the periphery (Fig. 11.1), although the molecular structure of the pigment is uniform for different compartments as well as maturation stages.

Although the astaxanthin spectra in the cell are nearly identical, they differ considerably from the Raman signature of the synthetic compound. Therefore, the study was complemented by quantum-chemical calculations [B3LYP/6-31+G(d,p)] that clarified that these differences were not related to the change of stereoisomeric form from 3*R*,3’*S* stereoisomer in the synthetic compound to predominantly 3*S*,3’*S* form in the case of algal cell [29].

As efficiency of astaxanthin production in the *Haematococcus* cells is, among others, related to temperature [37, 38], changes in Raman spectra of a single algal cell upon temperature treatment have been investigated and analyzed with the aid of quantum calculations at the B3LYP/6-31G(d,p) level [30]. Heating of a single *Haematococcus* cyst cell form −100°C to 150°C resulted in some spectral alterations that, based on the computations, were assigned to the conversion of the end-ring conformations in respect to the hydrocarbon chain. Particularly, for theoretically predicted Raman spectra of 3*S*,3’*S* conformers, (1) a decrease of the Raman shift of the bands due to the stretching vibration of the C=C bonds, (2) a decrease of the intensity of bands at *ca*. 1190 cm^{-1} relative to the bands at *ca*. 1160 cm^{-1}, with the increase of energy (and the end-ring conformations change from *gauche* to *trans*) were predicted. The effects opposite to changes 1 and 2 were observed upon heating of *Haematococcus* from −100°C systematically up to 150°C (Fig. 11.2).

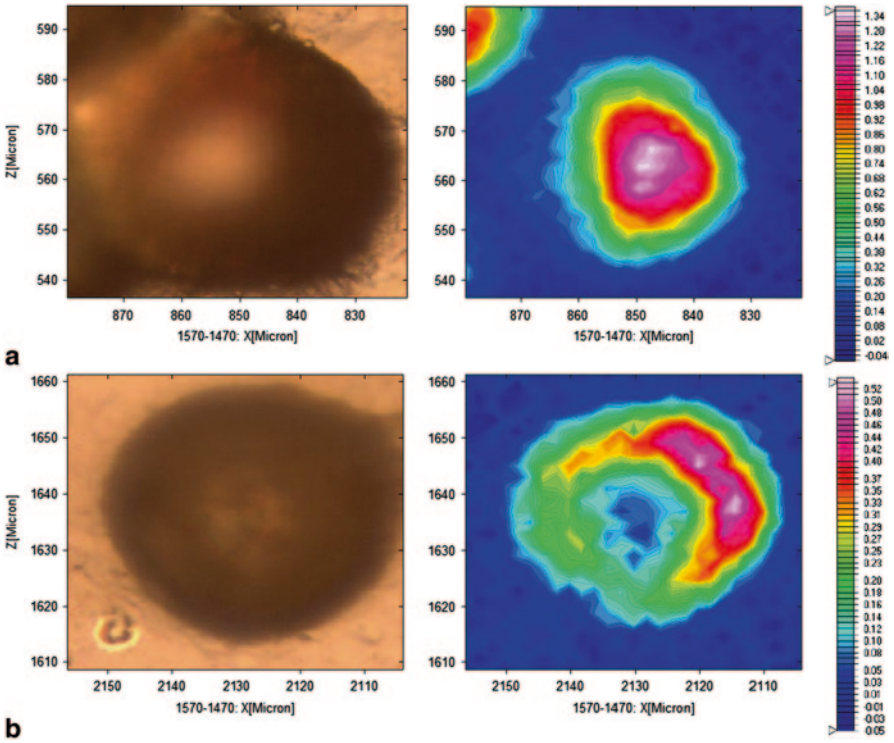


Fig. 11.1 A visual image (40×) focused inside of the spheroidal *Haematoccocus pluvialis* cell (**a**, on the *left*) and the cell with the collapsed cytoplasm (**b**, on the *left*) and the distribution of astaxanthin in the studied cells (on the *right*) obtained by integration of the marker band at 1520 cm^{-1} (the color code defined in the scales). The maps were obtained by measuring 900 and 729 points for **a** and **b**, respectively (the distance between points on the x and z axes equals to 2.0 μm) [29]. Reproduced from Ref. 29 with permission from The Royal Society of Chemistry

Therefore, it was proposed that the carotenoid, initially in the form of H-aggregates with the *trans* conformation of the both end-rings, is released from aggregates and rotamerizes to the more stable *gauche* forms upon the thermal treatment. Quantum-chemical calculations were very useful in predicting of the structural changes associated with heating [30].

Apart from implementation of quantum-chemical calculations as a complementary method to analyze experimental Raman data, other strategies are possible. A good approach toward distinguishing between different pigment molecules inside the cell and determining of the structure and distribution changes of carotenoids upon environmental stimuli is the application of statistical methods, such as multivariate analysis, to discern complicated experimental data.

Heraud et al. [13] applied chemometric methods to determine differences between individual living *Dunaliella tertiolecta* cells grown under nitrogen-replete and nitrogen-starved conditions. Two classification methods, namely soft independent modeling of class analogy (SIMCA) and discriminant partial least squares

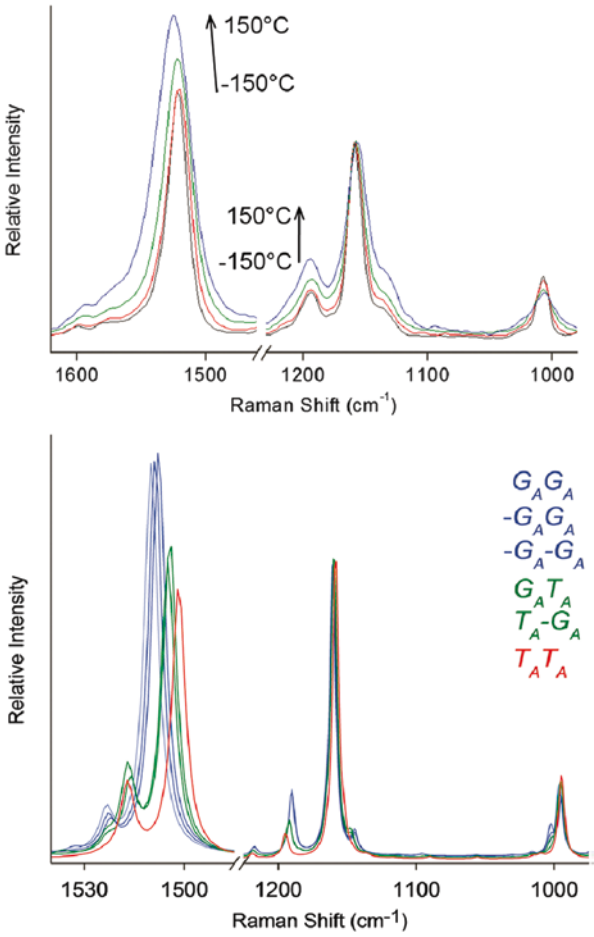


Fig. 11.2 FT-Raman spectra of *Haematococcus pluvialis* (*upper panel*) obtained during heating of the sample from −150 to 150 °C: *black*, *red*, *green* and *blue* line refers to −150, −50, 50, and 150 °C, respectively, compared with theoretical Raman spectra of AXT conformers (color coded, *lower panel*). All spectra were normalized by the height of the band at *ca*. 1157 cm^{-1}. (Reprinted (adapted) with permission from ref. [30]. Copyright (2013) American Chemical Society)

(PLS) were used to classify obtained Raman spectra of individual cells. With the applied pre-processing procedures, PLS resulted in much better overall classification than SIMCA, particularly when the second derivative using the Savitzky-Golay algorithm was applied. It was demonstrated that the last procedure was fully capable to extract chemical differences between cells grown in two mentioned nutrient conditions. As it is confirmed by the average spectra of nitrogen-replete and nitrogen-starved cells, the chlorophyll *a* bands are more intense in the nitrogen-replete population of cells compared to β-carotene bands, while the reverse effect is observed for nitrogen-starved cells [13]. Classification of the nitrogen-starved and replete *Dunaliella tertiolecta* cells was further improved by the analysis of mean spectra obtained from at least four individual spectra acquired at 2 μm intervals across the longitudinal axis of each cell (100 and 90 % of N-starved and N-replete cells, respectively, classified correctly) [14]. It was demonstrated that the analysis of first two PC loadings allowed for finding of the exact spectral features related to

differences between nitrogen-replete and starved cells, i. e. to elucidate chemical information encoded in the Raman spectra. It was stressed out that the single cell approach enables assessing population variability and the individual cellular answer for environmental stimuli, for instance nutrient conditions, without losing the overall information about the cellular system [14].

The insightful *in vivo* study of carotenogenesis in *Haematococcus Pluvialis* living cells by means of confocal Raman imaging supported by multivariate analysis was reported by Collins et al. [19] The application of Multivariate Curve Resolution (MCR) resulted in generating of the reconstructed cell images in which the relative concentration and location of four pure components (astaxanthin, β-carotene, chlorophyll and cellular autofluorescence) were determined at every spatial pixel (Fig. 11.3).

In flagellated motile cells, astaxanthin was practically absent, and only the weak intensity signal of this pigment was observed in the palmelloid cells. In both these morphotypes, β-carotene was found to be co-localized with chlorophyll in chloroplasts. Upon further stress conditions and production of palmelloid cells, distribution of both carotenoids changed and both co-localized β-carotene and astaxanthin were found outside the chloroplasts, supporting the hypothesis that upon stress conditions β-carotene is transported *via* a chloroplast membrane and transformed into astaxanthin. While β-carotene was found also in chloroplasts, astaxanthin is localized solely outside chloroplasts, in the center of the palmelloid cells. Finally, a massive production of astaxanthin was observed in aplanospores with the increased concentration of the pigment in the periphery of cysts. Therefore, the application of MCR analysis allowed not only for discrimination of two components of a very similar spectral profile (β-carotene and astaxanthin), but also for extracting of information about relatively weak spectral features (β-carotene bands are about 4-fold less intense compared to astaxanthin ones) [19].

Abbas et al. [15] described the application of target orthogonal partial least squares (T-OPLS) to the analysis of the subtle differences between carotenoid spectra of two species *Dunaliella* and *Phaeodactylum*. In T-OPLS, a known reference spectrum (in this case a standard carotenoid spectrum) is designated as the target and denoted as the single y column in an OPLS regression model, while the X matrix consists of the unfolded image spectra as variables in its columns. Finally, the carotenoids correlation images from *Dunaliella* and *Phaeodactylum* images in the model were reconstructed from the OPLS p loadings.

Both the original and reconstructed images showed distribution of carotenoids in the studied species. In *Dunaliella*, pigments are dispersed over the cell, with the high concentration of carotenoids in the eye, while in *Phaeodactylum* carotenoids are localized mainly in the center of cells. The orthogonal scores of the T-OPLS model enabled to find small differences between the reference and the studied samples, including variations between different samples of the same species (most probably due to the environmental influence) [15]. The work confirms the high potential of multivariate analysis in the extraction of detailed and exact information from the samples *in situ* in their complicated biological context.

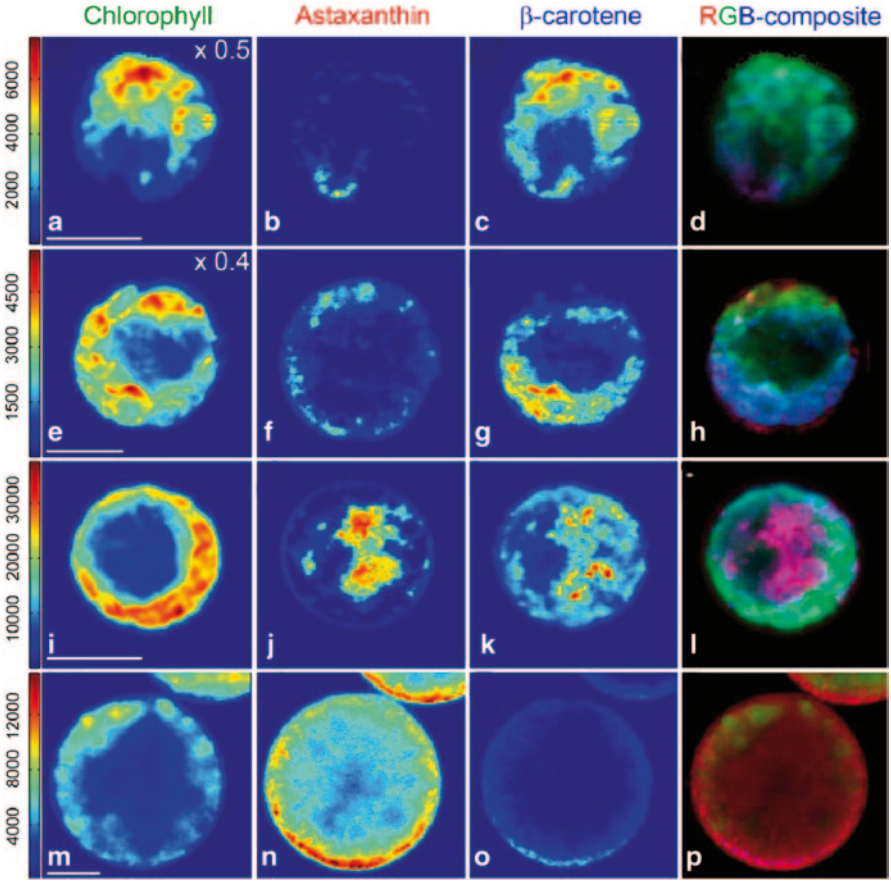


Fig. 11.3 MCR concentration maps for spectral components representing chlorophyll, astaxanthin, and β-carotene. **a–d**. Flagellated motile cell. **e–h**. Palmella stage cell under non-inductive conditions **i–l**. Palmelloid cell under inductive conditions for 24 h and **m–p**. Large red cyst (aplanospore). In all images the scale bar represents 10 mm. Note that in **a** and **e**, the intensity has been scaled by 0.5 and 0.4, respectively for clarity. The intensity scales are not comparable across cell types as the acquisition parameters were optimized for each cell image however, within a cell type the intensities within the concentration map represent relative component concentrations. Composite RGB images are created by overlaying the individual concentration maps that have been pseudo-colored as indicated in the image titles [19]. (Figure reprinted upon Creative Commons Attribution License (CCAL))

11.3 Carotenoids in Higher Plant and Single Animal Cells

Apart from the fact that β-carotene is found commonly in higher plants, it is also accumulated in various animal cells. Raman spectroscopy was used to investigate of this carotenoid both in plant [39, 40] and animal cells of several different systems

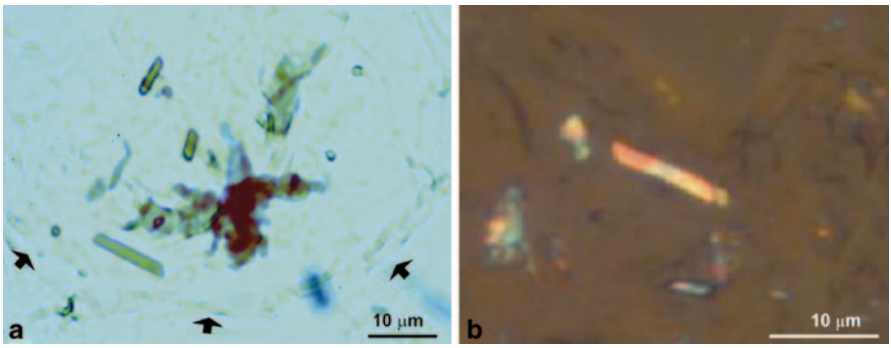


Fig. 11.4 Carotenoid crystals visible in a carrot cell under light microscope (**a**) and RamanScopeIII microscope (**b**). Arrows indicate the cell wall. (Reprinted from: In situ detection of a single carotenoid crystal in a plant cell using Raman microspectroscopy, Ref. [39] with permission from Elsevier)

e.g. in bovine corpus luteum cells [41, 42] and oocytes of frogs [43]. This section is a brief overview of these purely experimental works.

Raman microspectroscopy was successfully applied to the analysis of subcellular structures in the carrot root (Fig. 11.4) [39].

Single point FT-Raman spectra of orange crystals observed under 40x objective in carrot cells were measured in order to determine their composition [39]. It was confirmed that the studied specimens contain carotenoids, presumably β-carotene. As the wavenumber of the carotene C=C stretching marker band in the crystals differed from the wavenumber of this band in pure isolated β-carotene (1520 and 1515 cm^{-1} for carrot root and pure carotene, respectively), it was concluded that β-carotene in the crystal is bounded to protein and/or the crystals contain other carotenoids such as α-carotene ($\nu_{C=C}$ at 1521 cm^{-1}) and lutein ($\nu_{C=C}$ at 1522 cm^{-1}) [39].

Raman spectroscopy combined with Cryo Scanning Electron Microscopy (Cryo-SEM) was applied to determine distribution of carotene compounds in tomato and carrot emulsions containing 5% of oil [40]. Distribution of carotenoids in the cells of vegetables was studied after appropriate preparations including skin removal, heating and blending of samples. The structure of carrot/tomato cells and oil droplets was visualised by SEM and followed by determination of their chemical composition by means of Raman spectroscopy. The results showed that the Raman spectrum of the carrot/tomato cell material was very similar to the Raman spectrum of olive oil droplets. Both spectra were dominated by carotenoid signals indicating that carotenoids were dissolved and passed to the oil phase. It was concluded that the presence of oil is a crucial factor responsible for the increased bioaccessibility of carotenoids [40].

Localization of β-carotene in living bovine corpus luteum cells was determined with the help of Raman spectroscopy. It was previously demonstrated that the bovine corpus luteum has the highest concentration of β-carotene relative to other tissues [41]. The spectra of different types of subcellar structures revealed that

β-carotene is present in all organelles of corpus luteum. As expected, the highest concentration of β-carotene was found in lipid fractions due to the fact that carotenoids are soluble in the lipid droplets in the cell, while nuclei and mitochondria, i. e. hydrophilic cellular compartments, are regions of the lower concentration of carotenoids [41].

Raman imaging was also applied to obtain information on the cellular distribution of carotenoids in *Xenopus laevis* oocytes [43]. Oocytes of stage one were chosen as the object of study. Oogenesis for *Xenopus laevis* can be divided into six steps with the stage one being particularly convenient for spectroscopic analysis due to the characteristic transparent cytoplasm in the cells [44]. Based on distribution maps it was demonstrated that the highest concentration of β-carotene corresponds to cytoplasm and cellular membrane while the lowest carotenoid concentration is found in nucleus [43].

11.4 Carotenoids in Single Human Cells

The growing attention is focused on studying carotenoids in human cells. Research in this area are mainly related to lymphocytes due to the protective role played by carotenoids in the human immune system. In addition to studies devoted to distribution of carotenoids in lymphocytes [7–9], the effect of age [45] and health condition [10] of the organism was investigated by Raman spectroscopy and Raman imaging. Furthermore, Raman spectroscopy was used to assign the spectral signature of different type of cells: human colon adenocarcinoma [4], gastric carcinoma [5] and human breast cancer cells [6, 46, 47]. The possibility of using Raman spectroscopy as a diagnostic tool allowing for distinguishing of cancerous/malignant and healthy cells was also presented [5, 6, 46]. In order to analyse complex spectral data, chemometric methods were frequently applied in the works described in this section.

11.4.1 Lymphocytes

Leukocytes are the most diverse group of blood cells. They can be divided into following subclasses: granulocytes (neutrophils, basophils and eosinophils) and mononuclear cells (lymphocytes and monocytes) [7]. Different types of leukocytes can be distinguished due to the shape of nucleus, size and cellular components. Raman spectroscopy has been often used to determine the chemical composition of single blood cells based on some distinctive Raman bands [7–10, 41, 45, 48, 49].

The four selected types of leukocytes: lymphocytes, monocytes, eosinophils and neutrophils were measured with the application of Raman microimaging. The spectra of nuclei and cytoplasm compartments of studied leukocytes are shown in Fig. 11.5. [7] Based on Fig. 11.5, it can be noticed that only cytoplasm of lymphocytes contains carotenoids as confirmed by the presence of two characteristic bands at 1522 and 1158 cm^{-1}. No traces of carotenoids were observed for other types of

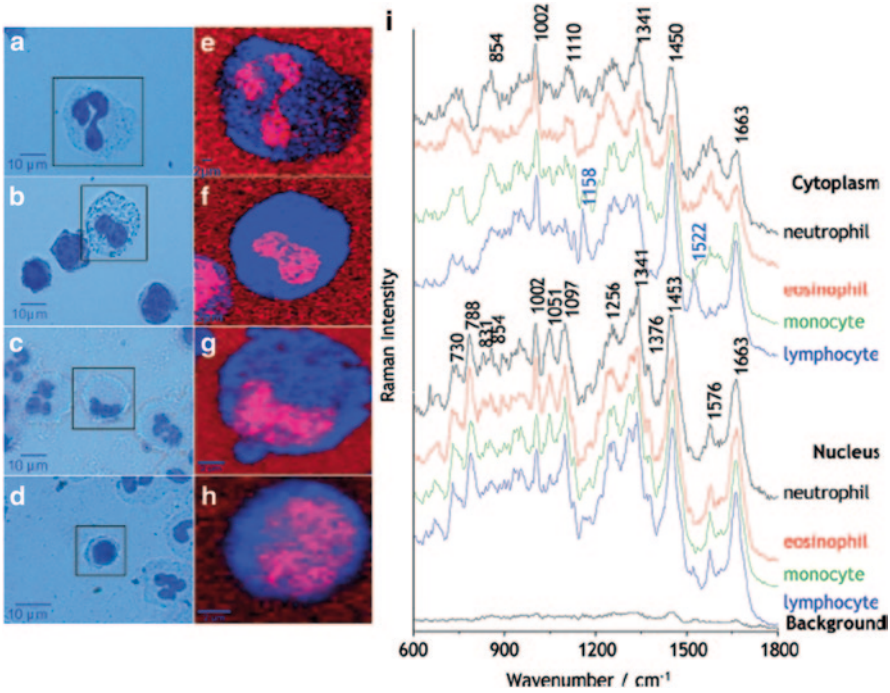


Fig. 11.5 Characteristics of leukocytes: a visual image after Lemura staining and false color Raman images of neutrophil (**a**, **e**), eosinophil (**b**, **f**), monocyte (**c**, **g**) and lymphocyte (**d**, **h**). Pink and blue areas indicate nuclei and cytoplasm, respectively. Raman distribution maps were obtained by integrating of the bands at 788 cm^{-1} and 2800–3050 cm^{-1}, for nuclei and cytoplasm, respectively. The respective average spectra of nuclei and cytoplasm are presented on the right. (Reprinted (adapted) with permission from Ref. [7]. Copyright (2013) American Chemical Society)

leukocytes [7]. Research of this type show the potential of application of Raman spectroscopy in the area of medical diagnostics.

Human lymphocytes play an important role in the acquired (adaptive) part of immune system. They are highly active cells that continually generate reactive oxidative species (ROS) as a part of their normal cellular activity. Even though ROS are produced as a part of the killing mechanism (e.g. to oxidize pollutants and viruses), they overproduction can lead to tissue damage. Therefore, the oxidant:antioxidant balance, enabling normal cellular function and health, needs to be maintained [50]. The role of antioxidants, including carotenoids e.g. β-carotene or astaxanthin, is the protection of immune cells against oxidative stress and membrane damage [51, 52]. Several works on carotenoids in the immune system studied by means of Raman spectroscopy have been published up to date.

Both distribution [7, 8] and concentration [9] of carotenoids in single human lymphocytes were studied with the help of Raman spectroscopy. Different types of cells: B-lymphocytes, T-lymphocytes and natural killer (NK) cells were studied in order to obtain information about subcellular distribution of carotenoids.

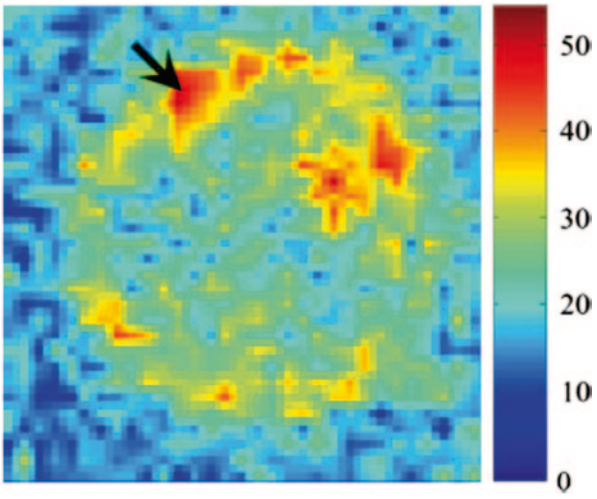


Fig. 11.6 Raman false image showing distribution of carotenoids in lymphocytes based on combination of bands 1154 and 1524 cm^{-1}. Red colour indicates the high concentration of carotenoids. (Reprinted (adapted) with permission from Ref. [8]. Copyright (2013) Wiley Online Library)

Additionally, the approximate concentration of carotenoids in single cells was estimated based on comparison of the Raman intensity of carotenoid signals in the solution (known concentration) and cells (unknown concentration) [9]. Based on the appearance of the marker bands at 1525 and 1157 cm^{-1} in the Raman spectra, the presence of carotenoids in the subcellular compartments was investigated. No characteristic bands of carotenoids in the Raman spectra of nuclei, neither in B- nor T-lymphocytes, nor NK-cells were found. On the contrary, the high concentration of carotenoids was observed in some places in cytoplasm for T-lymphocytes and NK-cells [9]. For both these types of cells, the particularly elevated concentration of carotenoids was found in Gall bodies (about 10^{-3} mol dm^{-3}) compared to the normal physiological carotenoids concentration in the blood plasma of 10^{-6} mol dm^{-3} [53]. As expected, the reason of the high concentration of carotenoids predominantly in the subcellular components containing high lipid concentration e.g. in Gall bodies is due to the lipophilic nature of these pigments [9]. Distribution of carotenoids in single cells was monitored using Raman imaging [8]. Raman distribution maps can be obtained from the sum of integral intensities of the marker bands at 1525 and 1157 cm^{-1}. In agreement with the former study [9], it was concluded that carotenoids are not homogeneously distributed in the cell, with their highest abundance in the cytoplasm and Gall bodies as shown in Fig. 11.6 [8].

It was additionally noticed that the amount of carotenoids in Gall bodies depended on the age and health condition of the organism. The intracellular carotenoids concentration in lymphocytes decreased significantly both with age [45] and in the case of lung cancer patients [10].

It was also demonstrated that diet is the factors affecting the amount of carotenoids in human lymphocytes, and that the supplementation of carotenoids e.g. addition of β-carotene to the diet leads to the increased concentration of carotenoids in lymphocytes. The influence of supplementation of 100 mg of β-carotene/day on the

carotenoid concentration in Gall bodies was analyzed in two advanced lung cancer patients and compared to the healthy subject. It was found that the carotenoids concentration increased during the days of supplementation but differently in the cytoplasm and Gall bodies. In the cytoplasm, the concentration of carotenoids raised after supplementation and continued to grow thereafter, but for Gall bodies the increase of the β-carotene concentration was observed only for first 5 days of supplementation, and remained constant thereafter. After supplementation the amount of β-carotene was the same for both lung cancer patients and the healthy subject in cytoplasm, while the amount of carotenoids in Gall bodies of lung cancer patients did not reach the concentration level of the healthy person. Therefore, it was concluded that carotenoids are not taken up to Gall bodies by simple diffusion [10].

11.4.2 Pathogenic Human Cells

Other human cells, where carotenoids were studied with the help of Raman spectroscopy, were human colon adenocarcinoma [4], gastric carcinoma [5] and human breast cancer cells [6, 46, 47]. It was suggested that in some cases carotenoids can be treated as Raman biomarkers to distinguish between the cancerous/malignant and healthy cells and tissues. The application of Raman spectroscopy and Raman imaging in medical diagnostic is a great hope for noninvasive and fast diagnosis in the future. Below-mentioned findings open the possibility of replacing the conventional histological imaging/staining techniques with high-resolution and highly sensitive Raman imaging.

In the study of human colon adenocarcinoma (cell line HT29), two groups of cells were treated with the medium supplemented with carotenoids: β-carotene or astaxanthin, respectively [4]. The aim of this study was to investigate the effect of carotenoids in the medium on the cells and verify the concentration of carotenoids after incubation. The concentration of carotenoids in the cells was estimated with the application of HPLC while pigments distribution was determined by Raman spectroscopy.

The cellular levels of carotenoids in living and metabolically active cells were measured after incubation with pigments for 24 and 72 h, respectively. For both carotenoids, β-carotene and astaxanthin, the highest cellular concentration of carotenoids without cytotoxic effects was observed with the concentration of 10 μmol dm^{-3} and the time of incubation of 72 h. The cellular level of astaxanthin in measured cells was approximately 50 times higher compared to β-carotene. For the chosen culture conditions, the distribution of carotenoids in colon cells was also determined. The highest intensity of the signal from astaxanthin (the distribution map obtained based on the band at 1525 cm^{-1}) was observed in the cytoplasm while the nucleus contained lower amount of astaxanthin. The Raman signal from β-carotene in cells was too low to generate Raman distribution maps of the latter pigment [4].

The findings for human colon adenocarcinoma cells coincide with the earlier studies referring to lymphocytes. The increased amount of carotenoids in specific places in the cytoplasm (for example Gall bodies) was observed in both types of

cells. In addition, no traces of carotenoids were found in the nuclei [7–10, 45]. As it was concluded for several other subjects, carotenoids as hydrophobic compounds prefer lipid-rich environment, therefore they accumulate in the cytoplasm containing less water than nuclei [4].

Raman spectroscopy was also used to characterize gastric carcinoma cells in both cultured conditions and in the human gastric mucosa tissue [5]. Authors aimed at determining of the differences between healthy and malignant tissues of human mucosa. In the first step, single Raman spectra of cultured cells were recorded. Further analysis based on the assignment of the bands showed that carotenoids are present in the cells, next to proteins, lipids, carbohydrates and nucleic acids. Furthermore, a series of single Raman spectra of normal and malignant gastric mucosa tissues was obtained and analyzed. Due to significant tissue heterogeneity, a number of small spectral variations inside each studied group i.e. the normal and malignant tissue was recognized. Nevertheless, small differences between the malignant and normal tissue were also found. The malignant tissue compared to the normal one is characterized by the decrease in the relative intensity of the bands at 1525 cm^{-1} (carotenoids), 1005 cm^{-1} (phenylalanine ring breathing) and 1660 cm^{-1} (amide I). Moreover, significant changes in the intensity ratio of the bands at 1525 and 1156 cm^{-1} were observed. The intensity ratio was higher for malignant gastric mucosa relative to the control that indicates for the increased content of carotenoids in the malignant tissue. This results demonstrate that carotenoids accumulation increased in the malignant gastric mucosa tissue relative to the control one [5].

Another potentially diagnostic application of Raman spectroscopy is the "optical biopsy" of breast cancer. Raman imaging was applied to characterize the breast cancer tissue and distinguish between noncancerous and cancerous specimens based on the spectroscopic features [6, 46, 47]. The most significant differences related to the cancer were found for marker bands due to carotenoids, lipids and proteins. Raman spectra of the cancerous tissue seem to be dominated by the protein component while the spectra of the noncancerous tissue are richer in lipids. Simultaneously, the carotenoids content is much lower in the case of the cancer tissue compared to the healthy one. These conclusions were obtained with the help of chemometrics by the application of cluster analysis (K-means algorithm). The presented results demonstrate that Raman spectra may be used to indicate cancerous changes in human breast [6, 46, 47].

11.5 Conclusions

The foregoing overview demonstrates that Raman spectroscopy and Raman imaging are sensitive and effective methods of studying carotenoids in single cells. Although research in this area focus on the experimental approach, first trials to use quantum-chemical methods as a tool to analyse Raman spectra of cells were already undertaken. Additionally, chemometric methods are used quite routinely

in the analysis of the complex Raman data of biological samples, including single cells.

Carotenoids were investigated in various types of cells, starting from unicellular organisms, single cells of higher plants and animals, to the individual human cells in the tissue environment. As described in the above examples, Raman spectroscopy can be used both for qualitative and quantitative analysis and, in the form of Raman imaging, is a particularly suitable method for studying distribution of carotenoids in single cells *in situ* and *in vivo*. In the case of unicellular organisms, carotenoids producers *de novo*, the impact of environmental conditions on the carotenoid distribution and structure has been also studied. Such studies are of high importance from the point of view of optimizing of the conditions of carotenoids production by unicellular organisms in the biofarms.

Higher organisms, including human beings supply carotenoids from the diet. With the help of Raman spectroscopy, it was demonstrated that the concentration of carotenoids in single cells of higher organism is changeable and depends on many factors. It was also demonstrated that in some cases the carotenoid concentration may be regarded as a biomarker of the organism disease state. This findings open the possibility to replace the conventional diagnostic methods by Raman spectroscopy and Raman imaging in the future.

References

1. Koyama Y (1995) Resonance Raman spectroscopy. In: Britton G (ed) Carotenoids Volume 1B: Spectroscopy, Birkhäuser Verlag AG, Basel, p 135–146
2. Ermakov IV, Sharifzadeh M, Ermakova M et al (2005) Resonance Raman detection of carotenoid antioxidants in living human tissue. J Biomed Opt 10:064028–064018
3. Gellermann W, Bernstein PS (2004) Noninvasive detection of macular pigments in the human eye. J Biomed Opt 9:75–85
4. Briviba K, Bornemann R, Lemmer U (2006) Visualization of astaxanthin localization in HT29 human colon adenocarcinoma cells by combined confocal resonance Raman and fluorescence microspectroscopy. Mol Nutr Food Res 50:991–995
5. Shen A, Ye Y, Zhang J et al (2005) Screening of gastric carcinoma cells in the human malignant gastric mucosa by confocal Raman microspectroscopy. Vib Spectrosc 37:225–231
6. Brozek-Pluska B, Musial J, Kordek R et al (2012) Raman spectroscopy and imaging: applications in human breast cancer diagnosis. Analyst 137:3773
7. Ramoji A, Neugebauer U, Bocklitz T et al (2012) Toward a spectroscopic hemogram: Raman spectroscopic differentiation of the two most abundant leukocytes from peripheral blood. Anal Chem 84:5335–5342
8. Pully V, Lenferink A, Otto C (2011) Time-lapse Raman imaging of single live lymphocytes. J Raman Spectrosc 42:167–173
9. Puppels G, Garritsen H, Kummer J et al (2005) Carotenoids located in human lymphocyte subpopulations and natural killer cells by Raman microspectroscopy. Cytometry 14:251–256
10. Bakker Schut TC, Puppels GJ, Kraan YM et al (1997) Intracellular carotenoid levels measured by Raman microspectroscopy: comparison of lymphocytes from lung cancer patients and healthy individuals. Int J Canc Prev 74:20–25
11. Majzner K, Kaczor A, Kachamakova-Trojanowska N et al (2013) 3D confocal Raman imaging of endothelial cells and vascular wall: perspectives in analytical spectroscopy of biomedical research. Analyst 138:603–610

12. Reddy RK, Bhargava R (2010) Chemometric Methods or Biomedical Raman Spectroscopy and Imaging. In: Matousek P (ed) Emerging Raman applications and techniques in biomedical and pharmaceutical fields. Springer Berlin Heidelberg, p 179–213
13. Heraud P, Wood BR, Beardall J et al (2006) Effects of pre-processing of Raman spectra on in vivo classification of nutrient status of microalgal cells. J Chemom 20:193–197
14. Heraud P, Beardall J, McNaughton D et al (2007) In vivo prediction of the nutrient status of individual microalgal cells using Raman microspectroscopy. FEMS Microbiol Lett 275:24–30
15. Abbas A, Josefson M, Abrahamsson K (2011) Characterization and mapping of carotenoids in the algae *Dunaliella* and *Phaeodactylum* using Raman and target orthogonal partial least squares. Chemometr Intell Lab 107:174–177
16. Milledge JJ (2011) Commercial application of microalgae other than as biofuels: a brief review. Rev Environ Sci Biotechnol 10:31–41
17. Brennan L, Owende P (2010) Biofuels from microalgae—a review of technologies for production, processing, and extractions of biofuels and co-products. Renev Sust Energ Rev 14:557–577
18. Spolaore P, Joannis-Cassan C, Duran E et al (2006) Commercial applications of microalgae. J Biosci Bioeng 101:87–96
19. Collins AM, Jones HD, Han D et al (2011) Carotenoid distribution in living cells of *Haematococcus pluvialis (Chlorophyceae)*. PLoS One 6:e24302
20. Tao Z, Wang G, Xu X et al (2011) Monitoring and rapid quantification of total carotenoids in Rhodotorula glutinis cells using laser tweezers Raman spectroscopy. FEMS Microbiol Lett 314:42–48
21. Pilat Z, Bernatova S, Jezek J et al (2012) Raman microspectroscopy of algal lipid bodies: beta-carotene quantification. J Appl Phycol 24:541–546
22. Han R-M, Zhang J-P, Skibsted LH (2012) Reaction Dynamics of Flavonoids and Carotenoids as Antioxidants. Molecules 17:2140–2160
23. Wormit M, Dreuw A (2007) Quantum chemical insights in energy dissipation and carotenoid radical cation formation in light harvesting complexes. Phys Chem Chem Phys 9:2917–2931
24. Wormit M, Dreuw A (2006) Carotenoid radical cation formation in LH2 of purple bacteria: a quantum chemical study. J Phys Chem 110:24200–24206
25. Walla PJ, Linden PA, Hsu C-P et al (2000) Femtosecond dynamics of the forbidden carotenoid S1 state in light-harvesting complexes of purple bacteria observed after two-photon excitation. Natl Acad Sci 97:10808–10813
26. Wang Y, Hu X (2002) A quantum chemistry study of binding carotenoids in the bacterial light-harvesting complexes. J Am Chem Soc 124:8445–8451
27. Neugebauer J, Veldstra J, Buda F (2011) Theoretical Spectroscopy of Astaxanthin in Crustacyanin Proteins: Absorption, Circular Dichroism, and Nuclear Magnetic Resonance. J Phys Chem B 115:3216
28. van Wijk AA, Spaans A, Uzunbajakava N et al (2005) Spectroscopy and quantum chemical modeling reveal a predominant contribution of excitonic interactions to the bathochromic shift in α-crustacyanin, the Blue carotenoprotein in the carapace of the lobster *Homarus gammarus*. J Am Chem Soc 127:1438–1445
29. Kaczor A, Turnau K, Baranska M (2011) In situ Raman imaging of astaxanthin in a single microalgal cell. Analyst 136:1109–1112
30. Kaczor A, Baranska M (2011) Structural Changes of Carotenoid Astaxanthin in a Single Algal Cell Monitored in Situ by Raman Spectroscopy. Anal Chem 83:7763–7770
31. Kubo Y, Ikeda T, Yang S-Y et al (2000) Orientation of carotenoid molecules in the eyespot of alga: *In Situ* polarized resonance Raman spectroscopy. Appl Spectrosc 54:1114–1119
32. Urban PL, Schmid T, Amantonico A et al (2011) Multidimensional analysis of single algal cells by integrating microspectroscopy with mass spectrometry. Anal Chem 83:1843–1849
33. Huang Y, Beal C, Cai W et al (2010) Micro-Raman spectroscopy of algae: composition analysis and fluorescence background behavior. Biotechnol Bioeng 105:889–898

34. He X, Allen J, Black P et al (2012) Coherent anti-Stokes Raman scattering and spontaneous Raman spectroscopy and microscopy of microalgae with nitrogen depletion. Biomed Opt Express 3:2896
35. Lu F-K, Ji M, Fu D et al (2012) Multicolor stimulated Raman scattering microscopy. Mol Phys 110:1927–1932
36. Boussiba S (2000) Carotenogenesis in the green alga Haematococcus pluvialis: cellular physiology and stress response. Physiol Plant 108:111–117
37. Fan L, Vonshak A, Boussiba S (1994) Effect of temperature and irradiance on growth of *Haematococcus Pluvialis (Chlorophyceae)*. J Phycol 30:829–833
38. Tjahjono AE, Hayama Y, Kakizono T et al (1994) Hyper-accumulation of astaxanthin in a green alga *Haematococcus pluvialis* at elevated temperatures. Biotechnol Lett 16:133–138
39. Baranska M, Baranski R, Grzebelus E et al (2011) In situ detection of a single carotenoid crystal in a plant cell using Raman microspectroscopy. Vib Spectrosc 56:166–169
40. Lopez-Sanchez P, Schumm S, Pudney PD et al (2011) Carotene location in processed food samples measured by Cryo In-SEM Raman. Analyst 136:3694–3697
41. Arikan Ş, Sands H, Rodway R et al (2002) Raman spectroscopy and imaging of β-carotene in live corpus luteum cells. Anim Reprod Sci 71:249–266
42. O'Fallon JV, Chew BP (1984) The subcellular distribution of β-carotene in bovine corpus luteum. In: Proceedings of the Society for Experimental Biology and Medicine Society for Experimental Biology and Medicine (New York, NY), vol 177, p 406–411
43. Rusciano G, Pesce G, Salemme M et al (2010) Raman spectroscopy of *Xenopus laevis* oocytes. Methods 51:27–36
44. Dumont JN (1972) Oogenesis in *Xenopus laevis (Daudin)*. I. Stages of oocyte development in laboratory maintained animals. J Morphol 136:153–179
45. Ramanauskaite RB, Segers-Nolten I, Grauw Kd et al (1997) Carotenoid levels in human lymphocytes, measured by Raman microspectroscopy. Pure Appl Chem 69:2131–2134
46. Abramczyk H, Brozek-Pluska B, Surmacki J et al (2011) Raman 'optical biopsy'of human breast cancer. Prog Biophys Mol Biol 108:74–81
47. Abramczyk H, Brozek-Pluska B, Surmacki J et al (2011) The label-free Raman imaging of human breast cancer. J Mol Liq 164:123–131
48. Puppels G, Garritsen H, Segers-Nolten G et al (1991) Raman microspectroscopic approach to the study of human granulocytes. Biophys J 60:1046–1056
49. Bankapur A, Zachariah E, Chidangil S et al (2010) Raman tweezers spectroscopy of live, single red and white blood cells. PLoS One 5:e10427
50. Chew BP, Park JS (2004) Carotenoid action on the immune response. J Nutr 134:257S–261S
51. Bolin AP, Macedo RC, Marin DP et al (2010) Astaxanthin prevents in vitro auto-oxidative injury in human lymphocytes. Cell Biol Toxicol 26:457–467
52. Jyonouchi H, Hill RJ, Tomita Y et al (1991) Studies of immunomodulating actions of carotenoids. 1. Effects of b-carotene and astaxanthin on murine lymphocyte functions and cell surface marker expression in vitro culture system. Nutr Cancer 16:93–105
53. Stacewicz-Sapuntzakis M, Bowen P, Kikendall J et al (1987) Simultaneous determination of serum retinol and various carotenoids: their distribution in middle-aged men and women. J Micronutr Anal 3:27–45

Part IV
Biomedical Applications

Chapter 12
Raman Spectra of Solid Amino Acids: Spectral Correlation Analysis as the First Step Towards Identification by Raman Spectroscopy

Tomasz Roliński, Sylwester Gawinkowski, Agnieszka Kamińska and Jacek Waluk

Abstract Raman spectra are presented for twenty proteinogenic amino acids, recorded for solid samples using four different laser wavelengths: 325, 514.5, 632.8, and 785 nm. In general, the intensity patterns are very similar for all excitation wavelengths. The spectra are analysed in terms of similarity. Possibilities of determination of amino acids in multicomponent powder mixtures are discussed and illustrated by examples.

Keywords Amino Acids • Raman Spectra • Correlation Analysis

12.1 Introduction

The amino acids are building blocks of proteins, hence the exact characterization of their properties is of fundamental importance in chemistry, biology, and medical sciences. Among various techniques used for that purpose, Raman spectroscopy is particularly attractive, due to its rich information content and sensitivity to such effects, as, e.g., conformational changes. Rapid progress in using Raman for the analysis of amino acids has been stimulated by the discovery of surface-enhanced Raman scattering (SERS) [1, 2], a technique which dramatically increases the signal intensity, and thus removes the main drawback of classical Raman methodology.

Detailed understanding of the vibrational structure of amino acids is a prerequisite for the investigation of polypeptides and proteins. However, from the experimental point of view this is not an easy task. In spite of numerous papers devoted to the subject, no database exists which would include the spectra of all important amino acids, measured under the same, well-defined and reproducible conditions. A recent step in this direction was provided by the work which compared the Raman spectra of solid amino acids with those measured in aqueous solutions [3]. Since the

J. Waluk (✉) · T. Roliński · S. Gawinkowski · A. Kamińska
Institute of Physical Chemistry, Polish Academy of Sciences, Kasprzaka 44/52, 01-224, Warsaw, Poland
e-mail: jwaluk@ichf.edu.pl

M. Baranska (ed.), *Optical Spectroscopy and Computational Methods in Biology and Medicine,* Challenges and Advances in Computational Chemistry and Physics 14,
DOI 10.1007/978-94-007-7832-0_12, © Springer Science+Business Media Dordrecht 2014

solution spectra are rather broad and noisy, the information content is definitively larger for the spectra of solids, which are much sharper, more intense, and noise-free. It is well known that in solution the amino acids can exist in different conformational forms, which makes the Raman spectra very difficult to interpret. Similarly to the situation in solution, amino acids in the solid phase are present as zwitterionic structures. Now, however, only one conformer is present, of which the structure is usually known from X-ray studies. Thus, by measuring the Raman spectra in solids, one avoids the problem of disentangling the mixture of conformers. Actually, different polymorphs are possible, which can be interconverted, e.g., by applying high pressure [4]. However, this complication should be considered an advantage rather than a drawback, since interconversion between different forms, if performed in a controlled fashion, can provide additional information on the vibrational structure.

In this work, we report the spectra of twenty basic amino acids, recorded for powder samples on the same instrument using four different laser wavelengths, ranging from UV (325 nm), through visible (514.5 and 632.8 nm) to the NIR (785 nm) region. We critically compare the spectra of the amino acids, using similarity analysis techniques. The connection with the possibility to identify particular amino acids in multi-component mixtures is underlined. Finally, examples of identification of components in mixtures containing up to eight amino acids are presented.

12.2 Materials and Methods

Samples The solid amino acids used for recording the Raman spectra were purchased from Aldrich and Merck. The samples from Aldrich included Glycine, *410225*; L-Alanine, *A7627*; L-Isoleucine, *I7403*; L-Asparagine anhydrous, *11149*; L-Cysteine, *168149*; L-Phenylalanine, *P212*; L-Tyrosine, *T8566*; L-Tryptophan, *T8941*; L-Serine, *S4500*; L-Glutamine, *G8540*. The compounds from Merck included L-Lysine Monohydrate, *1.12233.0100*; L-Valine, *1.08495.0025*; L-Leucine, *1.05360.0025*; L-Serine, *1.07769.0010*; L-Arginine, *1.01542.0100*; L-Aspartic acid, *1.00126.0100*; L-Proline, *1.07434.0010*; L-Glutamic acid, *1.00291.0250*; L-Methionine, *1.05707.0100*; L-Histidine, *1.04351.0025*; L-Threonine, *1.08411.0010*.

Measurements The spectra were recorded with Renishaw InVia Raman microscope, using four laser lines for each amino acid sample: 785 nm (Renishaw HPNIR785), 632.8 nm (HeNe, Renishaw RL633), 514.5 nm (Ar^+, Stellar Pro Modu-Laser, LLC) and 325 nm (HeCd, Kimmon). The spectrometer was equipped with 1200, 1800 and 2400 grooves/mm gratings, cutoff optical filters, 1024 × 256 pixels Peltier-cooled RenCam CCD detector, and an automatic translation stage. The 20× objective was used for visible and the 15× objective for UV excitation. The laser power on the sample was 50 mW or less. The spectral resolution was 5–6 cm^{-1} in the visible region and 10 cm^{-1} in the UV range, with the wavenumber accuracy of 2 cm^{-1}. The Raman shift was calibrated using both the Rayleigh line and the 520.6 cm^{-1} line of crystalline silicon. The wavelength dependence of

the spectrometer sensitivity was calibrated with a Bentham CL2 spectral irradiance standard lamp.

To prevent orientational dependence of Raman signals on the polarization of the laser beam, the crystalline samples were carefully ground and dispersed on microscopic glasses. At least 100 point maps of Raman spectra were recorded for each sample and then averaged. This procedure was repeated for all four laser lines. To prevent the influence of long term laser fluctuations on the relative intensities of all amino acids spectra, the laser power on the sample was monitored by measuring the intensity of the 520.6 cm^{-1} band of crystalline silicon.

Basics of statistical analysis The simplest method of spectral analysis is peak identification directly from a graph. We identify peaks by taking their positions and intensities and checking in a related database if they fit to some peaks of the reference spectrum of a substance. If the number of substances in the database increases, we take into account more peaks to make the method more specific. The next step is to consider spectra as continuous functions and look in the database for a spectrum most similar to the analysed one. As a standard measure of comparison for this task, one can choose the Pearson correlation coefficient, the Euclidean distance or the angle (dot product) between the spectra. It is obvious that this kind of analysis is well suited for computers, because doing it manually can be very tedious, especially when the number of substances in a database, together with the number of peaks to be considered, is growing. As a consequence, it is of no surprise that computers have been used in this domain for a long time. For instance, in [5] the authors identified single IR spectra by the Pearson correlation coefficient. They varied the wavenumber range and the number of data points to find the optimum values for this task. They also checked the stability of the method by shifting the wavenumber scale and by perturbing the intensities. These steps are necessary to account for the finite accuracy and reproducibility of the spectrometer, as well as for a possible presence of impurities in the sample.

In case of the spectrum of a mixture the identification is more complicated, even for non-interacting components, because in this simplest, ideal case the intensities are sums of component intensities. One may hope that for some spectral regions the intensities are due mostly to one component, i.e., this component is dominant in the mixture spectrum for a particular spectral interval. In an identification algorithm, one might look for a component spectrum most similar to the mixture spectrum, then subtract the component spectrum and search for successive components by comparing their spectra with the difference spectrum. This procedure stops when some kind of stop condition is satisfied. Several relevant parameters must be defined, such as the tolerance with which we take the peak positions and intensities. If the algorithm is based on the identification of peaks, one must decide how many peaks should be taken into account for the mixture and component spectra. In the case of treating spectra as continuous functions, one must define the width of a wavenumber interval to consider. The problem of appropriate interval choice for matching spectra with a correlation function was addressed in [6]. The authors presented a detailed analysis of simulated spectra. They varied peak positions and

intensities and concluded that similarity functions, such as correlation coefficient and Euclidean cosine, together with their first derivative variations should be used in windows where their values are more pronounced, which means that the differences in spectra will be highlighted. In order to decide what the values of the above parameters should be, one must devise an algorithm for a specific case, implement it on a computer, and select the best values by trial and error method.

Spectral correlation analysis can be considered as the preprocessing step for the identification of substances in the mixture of several components taken from the collection of substances that we are concerned with, by mixture spectrum and template spectra of substances. If we assume the linear model for a mixture spectrum, i.e., it is a linear combination of template spectra, the identification is an optimization process where we look for optimum component coefficients that minimize or maximize a comparison measure. In this situation it is evident that the more similar the spectra, the more difficult is their identification. Thus, it is important to perform this correlation analysis for different correlation measures and different parameters mentioned above in order to select the most appropriate ones. The values of the measures for given parameters and all pairs of template spectra can help in the determination of estimates from below (above), i.e., minimum (maximum) values for which we can accept a potential component as part of a mixture. The estimation from below (above) is related to the similarity (dissimilarity) measure. Let us remind here that the Pearson correlation coefficient is a similarity measure, whereas the Euclidean distance or the Euclidean cosine are dissimilarity measures.

Thus, if the value of a similarity function for a template and mixture spectra is substantially greater than its maximum value for all pairs of template spectra from the database, then we have a strong indication for the template to be part of the mixture. We can relax this condition by taking the maximum value of similarity function for this particular template and the rest of the spectra from the database. For the template spectrum that can be easily differentiated from the rest (e.g., Phenylalanine, *vide infra*) the condition is considerably weaker.

12.3 Results

We performed spectral similarity investigation for 20 most important amino acids coded by living organisms. They are listed in Table 12.1. As templates, we used the spectra recorded separately for each amino acid, using the laser line of 632.8 nm. The experimental spectra are presented in Figs. 12.1 and 12.2. Changing the laser wavelength does not have much influence on the spectra with respect to both, band positions and intensities. The amino acids containing the phenyl substituent exhibit absorption at the UV line; because of fluorescence, it was not possible to obtain the spectra for tryptophan and tyrosine for 325 nm excitation. The spectra of phenylalanine could be registered, but they were strongly overlapping with the emission.

Comparison of Raman spectra of a large group of proteinogenic amino acids in the solid phase has been, until now, presented in two papers. A work by Zhu et al.

Table 12.1 Amino acids of which the spectra were used in the similarity investigation

	Amino acid	Acronym
1	Arginine	Arg
2	Proline	Pro
3	Alanine	Ala
4	Phenylalanine	Phe
5	Cysteine	Cys
6	Asparagine	Asn
7	Glutamine	Gln
8	Leucine	Leu
9	Threonine	Thr
10	Valine	Val
11	Isoleucine	Ile
12	Glutamic acid	Glu
13	Glycine	Gly
14	Aspartic acid	Asp
15	Lysine	Lys
16	Methionine	Met
17	Histidine	His
18	Serine	Ser
19	Tryptophan	Trp
20	Tyrosine	Tyr

[3] contains the data for 18 amino acids. With respect to the data presented here, cysteine, glutamine, and asparagine are missing. However, these data lack the low frequency region, starting from values different for each amino acid, in the range of 400–670 cm^{-1}. The other paper, by Culka et al. [7], is focused on registering the spectra outdoors, using two different portable Raman instruments. This study is limited to 13 amino acids. Below we present a concise summary of the relevant Raman spectra of solid amino acids registered so far, noting, for several cases, the discrepancies between these spectra and our results.

Glycine Many groups studied the Raman spectra of the simplest amino acid in the solid phase [3, 7–14]. The spectra reported by various researchers are quite similar. The Raman spectrum of solid glycine presented in the most recent work [3] practically coincides (within 2–3 cm^{-1}) with our spectra. Also the relative intensity patterns are comparable.

Alanine This molecule is amongst the most popular of all amino acids, and therefore numerous workers used Raman techniques in the investigation of the crystalline form, including pressure and temperature dependence of the spectra [3, 7, 15–27]. Our spectrum nicely agrees with those recorded for both polycrystalline and single crystal samples.

Arginine In contrast to alanine, literature data on arginine are rather scarce. The Raman spectra of several salts have been reported [28–32]; they strongly differ from a spectrum of solid arginine published by Zhu et al [3]. The latter strongly resembles our spectrum.

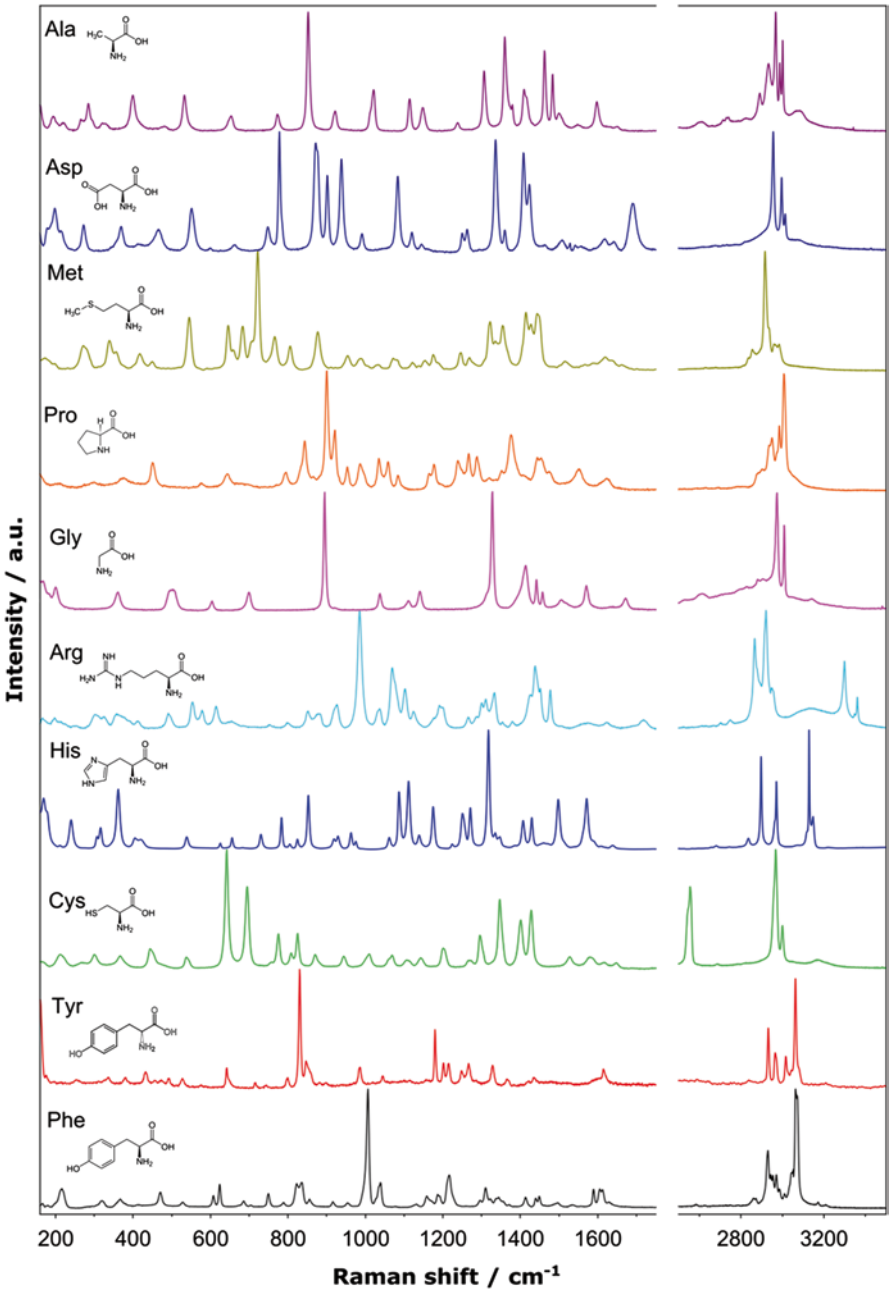


Fig. 12.1 Experimental spectra of solid amino acids measured at 293 K using 632.8 nm laser line

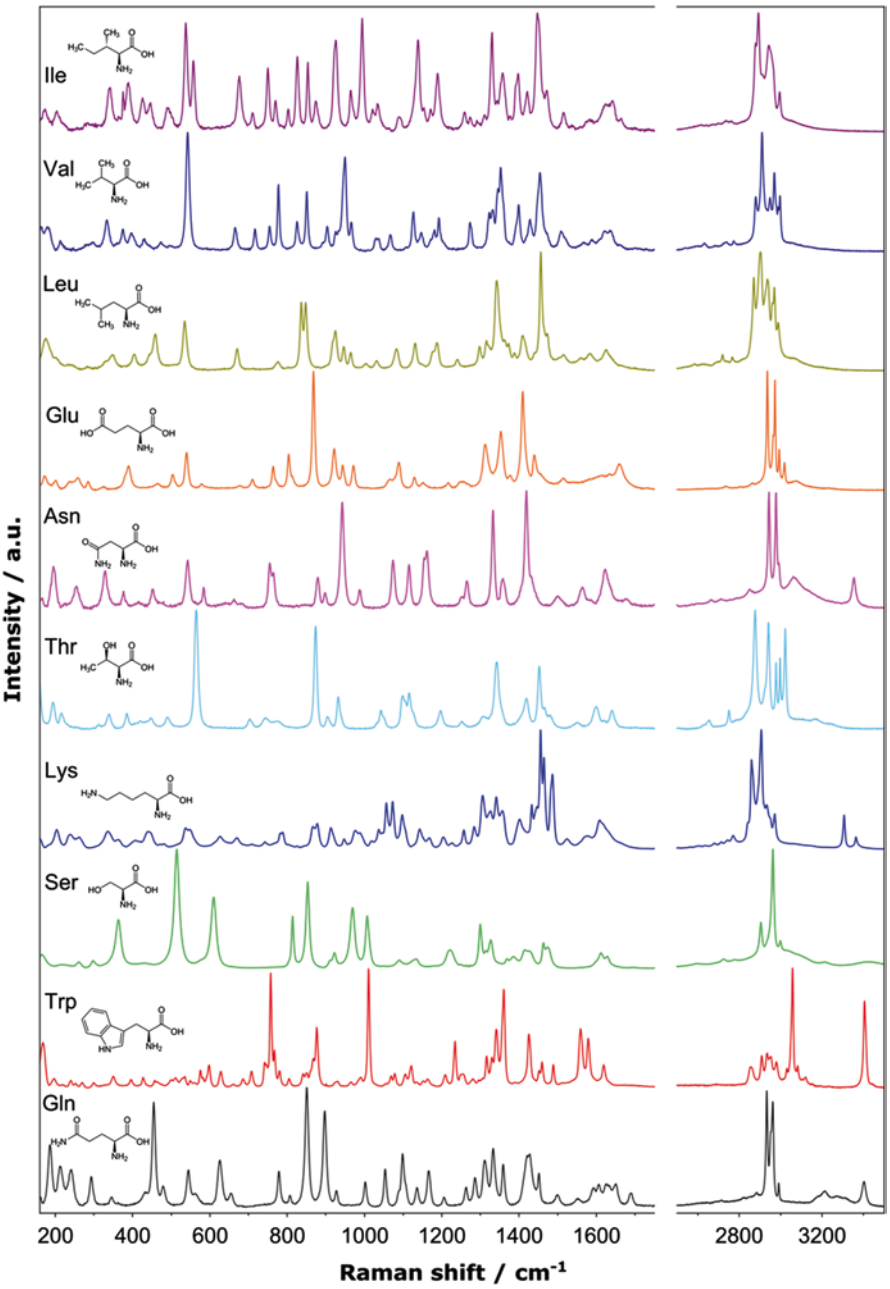


Fig. 12.2 Experimental spectra of solid amino acids measured at 293 K using 632.8 nm laser line

Asparagine We found only three references to the works describing Raman spectrum of solid asparagine [7, 33, 34]. The frequencies and intensities reported recently [7] agree well with our spectrum.

Aspartic acid There are only a few papers related to the Raman spectra of aspartic acid [3, 7, 35]. The published data agree well with our spectra regarding peak positions, but the relative intensities, even among strong peaks, are significantly different.

Cysteine Raman spectra of solid cysteine were reported in several papers [36–39]. The frequencies in the Raman spectrum of solid L-cysteine, published in 1983 [37], coincide within 5 cm^{-1} with our spectra. The exceptions are (i) the peak reported at 931 cm^{-1}, corresponding to our value of 944 cm^{-1}; (ii) a peak we observe at 1069 cm^{-1}, which was not mentioned. No intensities were reported in this work. In a recent paper [39], the frequency of 944 cm^{-1} was given, in exact agreement with our results; a peak corresponding to our 1069 cm^{-1} band is reported at 1064 cm^{-1}.

Glutamine Very few studies have been devoted to the Raman spectroscopy of solid glutamine [7, 40, 41]. Our data compare well with those reported by Culka et al. [7].

Glutamic acid Literature reports on the Raman spectroscopy of glutamic acid are limited to a few papers [3, 7, 42, 43]. Our spectra agree very well with those published by Culka et al. [7]. On the contrary, comparison with the spectra reported by Zhu [3] reveals significant differences in peak positions. Most probably, these authors used another crystalline form of this substance.

Isoleucine Our spectra agree well with those reported by Zhu et al. [3] and Almeida et al. [44] with regard to the observed frequencies. A band we observe at 925 cm^{-1} was reported by Zhu at 918 cm^{-1}; this band was observed by Almeida for L-isoleucine crystals at 922 or 924 cm^{-1}, depending on the scattering geometry. Interestingly, the Raman spectrum reported by Sabino et al. in a study of high pressure effects [45] is quite different with regard to both peak locations and intensities, even though the authors claim to have used the same crystal form as Almeida.

Leucine Raman spectra of solid leucine have been reported in a number of papers [3, 46–52]. Our spectra agree very well with those of Zhu et al. [3], although the relative intensities are somewhat different.

Histidine Not much work has been reported for solid histidine [3, 53]. Our data agree with those by Zhu et al. [3]. A Raman spectrum reported for $NaBF_4$—doped L-histidine [54] is completely different.

Lysine Our spectra agree nicely with those of Zhu [3]. Aliaga et al. [55] reported the Raman spectrum of solid lysine which resembles our spectrum in the high energy region, but is different in the range below 1000 cm^{-1}.

Methionine The published Raman spectra of solid methionine [3, 7, 56, 57] agree well with our spectra with respect to both frequencies and intensities.

Threonine Except for minor relative intensity variations, our spectra are very similar to those reported by Zhu [3], as well as to the data published by other workers [7, 58].

Phenylalanine Our spectrum agrees with the spectrum published previously for a somewhat smaller wavenumber range [3].

Proline Quite a few papers discussed the Raman spectra of solid proline [3, 7, 59–62]. Our data are in good agreement with previously published results regarding peak positions, but the relative intensities are significantly different from those reported by Zhu et al [3].

Serine Literature data on solid serine are quite extensive [3, 7, 37, 63–67]. Our spectra compare satisfactorily with those of Culka et al. [7], whereas the spectra reported by Zhu et al. [3] are quite different with regard to intensities.

Tryptophan Raman spectra of solid tryptophan have been reported in several papers [3, 7, 68–70]. Our spectra are in very good agreement with those published recently [3, 7].

Tyrosine Three recent papers reported the Raman spectra of solid tyrosine [3, 7, 71]. These data correspond well to our spectra. It is to be noted, though, that the relative intensities differ from those shown in the work by Zhu et al. [3]. In particular, in the region of 1100–1400 cm^{-1} the dominant peak is located at 1328 cm^{-1}, whereas in our spectra the most intense band is observed in this range at 1198 cm^{-1}.

Valine Different groups have studied Raman spectra of valine in the solid state [3, 52, 66, 72, 73]. Our spectra agree well with frequencies detected for the crystals, but the intensities follow a different pattern, very similar to that observed by Zhu et al [3]. Interestingly, in one case, for the spectra recorded in nujol mulls [52], the measured frequencies are completely different from those reported for all other solid samples.

It should be stressed that, due to polarization effects, the relative intensities of the band may change when the crystalline sample is pulverized. Our data refer to finely ground powders, for which these effects should be minimized. This may be the reason for discrepancies in relative intensities between our measurements and other works in the spectra discussed above.

12.3.1 Similarity Analysis Using Pearson Correlation Coefficient

We start the analysis using Pearson correlation coefficients between the template spectra of single amino acids in the range from 300 to 1700 cm^{-1}. The values of the correlation coefficients for all pairs are presented in Fig. 12.3. It is easy to see that they have a tendency to increase towards the lower-right part of the table. This is a consequence of sorting amino acids according to the mean value of the coefficient for a given template spectrum and the rest of the spectra. Let us also note that this

	Phe	Tyr	Cys	His	Arg	Gly	Pro	Met	Asn	Ala	Glu	Trp	Ser	Lys	Thr	Asp	Gln	Leu	Val	Ile
Phe	100	11	-2	-3	-4	-3	3	-8	-7	3	-6	11	21	-5	-9	-8	5	6	-3	5
Tyr	11	100	11	5	2	2	25	-1	0	5	-7	-4	13	-9	-9	-9	16	14	3	13
Cys	-2	11	100	-5	-1	-5	3	23	13	-1	16	9	-1	5	13	19	10	13	17	9
His	-3	5	-5	100	14	29	6	11	4	10	14	9	20	10	8	7	19	16	17	12
Arg	-4	2	-1	14	100	13	9	11	18	-1	14	8	19	26	14	10	17	8	15	20
Gly	-3	2	-5	29	13	100	16	15	17	-2	12	13	24	15	12	17	29	10	17	26
Pro	3	25	3	6	9	16	100	-3	-1	13	13	-1	13	26	11	19	30	41	16	18
Met	-8	-1	23	11	11	15	-3	100	21	3	21	19	2	15	20	23	13	16	34	33
Asn	-7	0	13	4	18	17	-1	21	100	5	23	31	5	8	26	45	28	13	35	24
Ala	3	5	-1	10	-1	-2	13	3	5	100	21	23	49	33	11	6	37	40	29	33
Glu	-6	-7	16	14	14	12	13	21	23	21	100	21	22	12	34	42	14	24	25	33
Trp	11	-4	9	9	8	13	-1	19	31	23	21	100	26	21	30	30	24	23	26	32
Ser	21	13	-1	20	19	24	13	2	5	49	22	26	100	18	13	11	42	29	21	29
Lys	-5	-9	5	10	26	15	26	15	8	33	12	21	18	100	33	23	20	48	45	36
Thr	-9	-9	13	8	14	12	11	20	26	11	34	30	13	33	100	57	26	29	27	36
Asp	-8	-9	19	7	10	17	19	23	45	6	42	30	11	23	57	100	29	24	33	25
Gln	5	16	10	19	17	29	30	13	28	37	14	24	42	20	26	29	100	37	29	32
Leu	6	14	13	16	8	10	41	16	13	40	24	23	29	48	29	24	37	100	60	48
Val	-3	3	17	17	15	17	16	34	35	29	25	26	21	45	27	33	29	60	100	65
Ile	5	13	9	12	20	26	18	33	24	33	33	32	29	36	36	25	32	48	65	100

Fig. 12.3 Values of the correlation coefficients (multiplied by 100) for all pairs of template spectra of single amino acids, calculated in the range from 300 to 1700 cm^{-1}. The tendency for the values to increase towards the lower-right corner of the *rectangle* can be observed here as a consequence of ordering of the successive amino acids with respect to the mean values of corresponding rows (or columns). The colours reflect subdivision of the values of the coefficient c into three essentially equal parts. They are defined by ranges $c \leq 11$, $11 < c \leq 23$, $23 < c$, which implies that the boundaries 11, 23 are 3-quantiles (tertiles)

table is symmetrical, which reflects symmetry of the correlation coefficient. The mean and standard deviation values for all coefficients in Fig. 12.3 are 0.17 and 0.14 respectively. It is natural to assume that similarities in the spectra will not be the same for different spectral regions. Therefore, as has already been suggested [6], the correlation coefficient should be calculated for selected subintervals rather than for the whole region. We thus repeated the same calculations for two equal intervals: from 300 to 1000 and from 1000 to 1700 cm^{-1} (see Fig. 12.4). For the lower frequency interval the mean and standard deviation are 0.07 and 0.16 respectively, whereas for the higher frequency region we obtain 0.27 and 0.20. The difference in values can also be seen by comparing visually the number of coefficients of the same color for the left and right panels of Fig. 12.4. To compare them also with Fig. 12.3 we retained the boundary values (11 and 23) that define the way they are color-coded. Obviously for Fig. 12.4 they are no longer 3-quantiles (tertiles).

	Phe	Tyr	His	Ala	Pro	Cys	Arg	Asn	Trp	Lys	Gly	Gln	Glu	Asp	Ser	Thr	Leu	Ile	Val	Met
Phe	100	3	-7	0	-2	4	-6	-8	17	-6	1	4	-1	-9	30	-4	1	2	-3	5
Tyr	3	100	13	-15	10	6	4	19	-1	-10	21	12	-7	-3	11	-5	-5	13	8	13
His	-7	13	100	-3	8	-1	14	5	9	3	42	20	20	13	18	14	15	7	16	29
Ala	0	-15	-3	100	17	4	-3	9	32	48	6	14	29	17	40	28	38	30	26	20
Pro	-2	10	8	17	100	1	18	-7	9	53	16	10	9	17	23	23	52	35	45	29
Cys	4	6	-1	4	1	100	31	22	40	15	4	23	41	40	25	42	26	28	40	31
Arg	-6	4	14	-3	18	31	100	32	10	25	29	37	33	25	30	29	18	26	32	47
Asn	-8	19	5	9	-7	22	32	100	32	3	33	57	34	49	29	38	12	30	22	45
Trp	17	-1	9	32	9	40	10	32	100	18	17	33	27	37	48	37	35	36	38	42
Lys	-6	-10	3	48	53	15	25	3	18	100	21	24	23	23	40	43	64	40	56	34
Gly	1	21	42	6	16	4	29	33	17	21	100	39	34	29	54	27	26	49	37	49
Gln	4	12	20	14	10	23	37	57	33	24	39	100	40	47	39	53	27	45	33	55
Glu	-1	-7	20	29	9	41	33	34	27	23	34	40	100	63	34	45	46	43	42	62
Asp	-9	-3	13	17	17	40	25	49	37	23	29	47	63	100	38	63	49	36	40	53
Ser	30	11	18	40	23	25	30	29	48	40	54	39	34	38	100	44	45	48	46	43
Thr	-4	-5	14	28	23	42	29	38	37	43	27	53	45	63	44	100	65	48	61	54
Leu	1	-5	15	38	52	26	18	12	35	64	26	27	46	49	45	65	100	62	79	57
Ile	2	13	7	30	35	28	26	30	36	40	49	45	43	36	48	48	62	100	85	72
Val	-3	8	16	26	45	40	32	22	38	56	37	33	42	40	46	61	79	85	100	69
Met	5	13	29	20	29	31	47	45	42	34	49	55	62	53	43	54	57	72	69	100

	Gly	Met	Phe	Arg	Cys	Trp	Thr	Tyr	Lys	Asn	His	Ser	Glu	Asp	Pro	Leu	Val	Ile	Ala	Gln
Gly	100	-9	-12	-8	-10	-1	-7	-9	-12	-7	-11	-1	-5	1	17	-13	-10	-14	-13	16
Met	-9	100	-22	-11	19	3	1	-6	2	5	-8	-18	2	3	-15	-14	13	4	-6	-10
Phe	-12	-22	100	-2	-8	-4	-19	21	-13	-6	9	13	-13	-8	7	14	-5	8	5	8
Arg	-8	-11	-2	100	-16	-2	-2	3	30	-1	12	11	2	-5	5	-5	-4	8	-1	-4
Cys	-10	19	-8	-16	100	-8	1	12	13	12	-9	-11	6	10	3	8	7	1	-1	7
Trp	-1	3	-4	-2	-8	100	20	-7	-8	26	-15	5	22	20	-9	-4	2	15	5	5
Thr	-7	1	-19	-2	1	20	100	-10	16	13	-10	-8	28	51	6	-10	-7	20	-7	1
Tyr	-9	-6	21	3	12	-7	-10	100	-14	-11	2	14	-8	-12	30	29	2	18	16	20
Lys	-12	2	-13	30	13	-8	16	-14	100	16	11	-14	5	35	15	7	25	10	-4	5
Asn	-7	5	-6	-1	12	26	13	-11	16	100	-4	-17	17	40	4	12	48	11	-3	-5
His	-11	-8	9	12	-9	-15	-10	2	11	-4	100	38	12	-8	10	19	19	21	43	19
Ser	-1	-18	13	11	-11	5	-8	14	-14	-17	38	100	16	-9	9	18	3	15	58	45
Glu	-5	2	-13	2	6	22	28	-8	5	17	12	16	100	28	14	7	13	28	16	-3
Asp	1	3	-8	-5	10	20	51	-12	35	40	-8	-9	28	100	21	-6	24	8	-7	11
Pro	17	-15	7	5	3	-9	6	30	15	4	10	9	14	21	100	39	1	9	11	42
Leu	-13	-14	14	-5	8	-4	-10	29	7	12	19	18	7	-6	39	100	36	26	41	46
Val	-10	13	-5	-4	7	2	-7	2	25	48	19	3	13	24	1	36	100	36	30	22
Ile	-14	4	8	8	1	15	20	18	10	11	21	15	28	8	9	26	36	100	37	14
Ala	-13	-6	5	-1	-1	5	-7	16	-4	-3	43	58	16	-7	11	41	30	37	100	57
Gln	16	-10	8	-4	7	5	1	20	5	-5	19	45	-3	11	42	46	22	14	57	100

Fig. 12.4 Values of the correlation coefficients (multiplied by 100) for all pairs of template spectra of single amino acids, calculated in the range from 300 to 1000 cm^{-1} (*right* panel) and from 1000 to 1700 cm^{-1} (*left* panel). For comparison reasons the ranges of *c* values are the same as in Fig. 12.3; however the boundaries are no longer quantiles

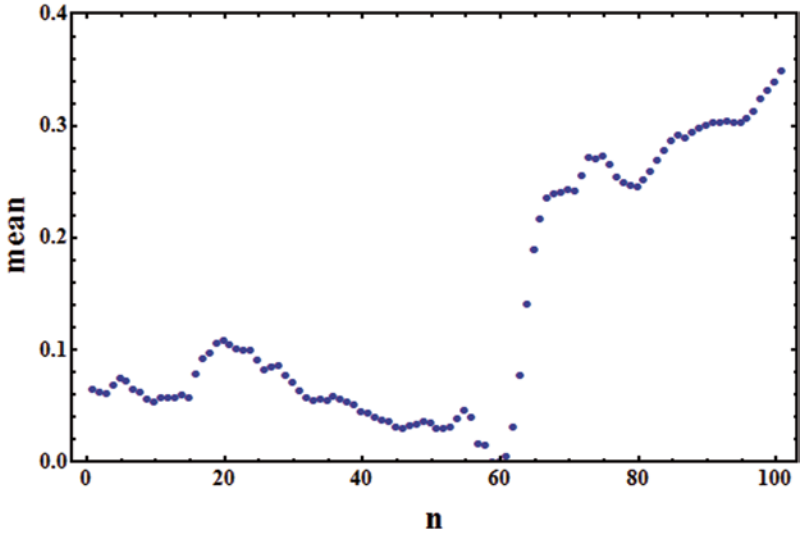


Fig. 12.5 The mean value of the correlation coefficient for all pairs of template spectra of single amino acids for a series of 400 cm^{-1} wide windows covering the whole range of wavenumbers considered (from 300 to 1700 cm^{-1}); n is the number of a succesive window, with lower and upper limits given by $300+(n-1)\cdot 10$ cm^{-1} and $700+(n-1)\cdot 10$ cm^{-1}, respectively, for n = 1, 2, …, 101

Since the low frequency interval is considerably better, i.e., the spectra are more distinguishable, it is of interest to further refine the structure of correlation for different parts of the spectra. Assuming that a 400 cm^{-1} interval can be used for this purpose, we define a series of windows of this width covering the whole interval from 300 cm^{-1} to 1700 cm^{-1} by shifting them in 10 cm^{-1} steps. For each window we calculated the mean of the correlation coefficients. The result is shown in Fig. 12.5. The minimum (0.00) of the mean correlation coefficient is attained for window number 59, which corresponds to the interval 880–1280 cm^{-1}. It can be seen that moving the window from this position towards higher frequencies results in a sharp increase in the mean value. The maximum (0.35) of the mean is attained for the rightmost window number 101 (see Fig. 12.5), which is the interval 1300–1700 cm^{-1}. Figure 12.6 shows all coefficients for these two extreme cases. The above analysis strongly suggests that it might be advantageous to perform the identification process for the component spectra of a mixture of amino acids in the range from 300 cm^{-1} to 1280 cm^{-1}, where the mean correlation coefficient and standard deviation equal 0.05 and 0.14, respectively.

12.3.2 Similarity Analysis Based on Separate Peaks

Instead of treating spectra as continuous functions one can perform analysis based on the identification of separate peaks. After recording peak positions and

	His	Ala	Tyr	Phe	Lys	Cys	Gly	Asn	Trp	Pro	Arg	Asp	Glu	Leu	Gln	Ile	Thr	Val	Ser	Met
His	100	-2	12	16	5	5	50	4	15	7	17	7	21	17	23	10	14	18	32	32
Ala	-2	100	-3	14	53	6	-2	2	29	41	5	14	31	41	19	27	24	25	37	14
Tyr	12	-3	100	39	0	0	49	45	16	4	28	9	5	-2	44	35	9	19	47	30
Phe	16	14	39	100	16	12	13	10	8	9	24	7	37	20	49	29	31	24	19	48
Lys	5	53	0	16	100	15	13	-2	17	69	27	16	15	70	18	38	46	57	51	28
Cys	5	6	0	12	15	100	6	34	52	13	33	56	50	34	35	37	57	47	30	40
Gly	50	-2	49	13	13	6	100	39	16	23	42	26	33	21	45	45	23	33	65	47
Asn	4	2	45	10	-2	34	39	100	42	3	43	56	37	14	76	39	45	26	50	46
Trp	15	29	16	8	17	52	16	42	100	22	26	40	30	40	47	43	48	45	32	45
Pro	7	41	4	9	69	13	23	3	22	100	33	29	21	71	21	55	49	66	58	41
Arg	17	5	28	24	27	33	42	43	26	33	100	35	35	25	54	43	37	39	64	64
Asp	7	14	9	7	16	56	26	56	40	29	35	100	61	47	56	35	73	41	46	50
Glu	21	31	5	37	15	50	33	37	30	21	35	61	100	46	47	45	50	40	38	62
Leu	17	41	-2	20	70	34	21	14	40	71	25	47	46	100	33	60	77	83	52	54
Gln	23	19	44	49	18	35	45	76	47	21	54	56	47	33	100	55	59	45	58	69
Ile	10	27	35	29	38	37	45	39	43	55	43	35	45	60	55	100	54	89	58	75
Thr	14	24	9	31	46	57	23	45	48	49	37	73	50	77	59	54	100	71	58	59
Val	18	25	19	24	57	47	33	26	45	66	39	41	40	83	45	89	71	100	59	69
Ser	32	37	47	19	51	30	65	50	32	58	64	46	38	52	58	58	58	59	100	51
Met	32	14	30	48	28	40	47	46	45	41	64	50	62	54	69	75	59	69	51	100

	Tyr	Phe	Lys	Trp	Arg	Ala	Met	Ser	Gly	His	Ile	Gln	Cys	Pro	Thr	Val	Asn	Leu	Asp	Glu
Tyr	100	-7	-17	-7	1	-25	3	-22	-11	13	-6	-15	15	-7	-18	-11	-11	-14	-17	-29
Phe	-7	100	-16	40	-14	2	-16	51	-7	-26	5	-8	5	-10	-19	-15	-17	-8	-19	-20
Lys	-17	-16	100	-22	35	-20	-9	-22	-2	-7	-9	12	-2	17	-4	-4	-6	-16	-9	21
Trp	-7	40	-22	100	-12	22	-7	66	-6	-11	-11	-5	1	-18	-11	-28	-11	-21	-14	-21
Arg	1	-14	35	-12	100	-16	7	-2	-15	4	3	-10	0	-10	-3	-6	-5	-9	-14	15
Ala	-25	2	-20	22	-16	100	-9	26	-2	0	15	-9	-11	-8	20	-17	4	0	-16	6
Met	3	-16	-9	-7	7	-9	100	-25	8	6	-10	-15	-8	-7	-10	18	27	9	16	-10
Ser	-22	51	-22	66	-2	26	-25	100	-11	-17	8	-15	-7	-19	-9	-7	-24	6	-17	12
Gly	-11	-7	-2	-6	-15	-2	8	-11	100	-16	-7	46	-13	30	3	-4	-4	-13	10	-10
His	13	-26	-7	-11	4	0	6	-17	-16	100	1	8	-19	2	16	-2	-5	15	11	21
Ile	-6	5	-9	-11	3	15	-10	8	-7	1	100	-3	-11	2	2	2	-21	38	-17	15
Gln	-15	-8	12	-5	-10	-9	-15	-15	46	8	-3	100	-12	21	28	-23	-1	-16	15	-3
Cys	15	5	-2	1	0	-11	-8	-7	-13	-19	-11	-12	100	-21	8	32	34	-5	16	8
Pro	-7	-10	17	-18	-10	-8	-7	-19	30	2	2	21	-21	100	0	3	-8	15	31	34
Thr	-18	-19	-4	-11	-3	20	-10	-9	3	16	2	28	8	0	100	2	19	1	25	9
Val	-11	-15	-4	-28	-6	-17	18	-7	-4	-2	2	-23	32	3	2	100	33	50	18	24
Asn	-11	-17	-6	-11	-5	4	27	-24	-4	-5	-21	-1	34	-8	19	33	100	11	49	21
Leu	-14	-8	-16	-21	-9	0	9	6	-13	15	38	-16	-5	15	1	50	11	100	11	48
Asp	-17	-19	-9	-14	-14	-16	16	-17	10	11	-17	15	16	31	25	18	49	11	100	29
Glu	-29	-20	21	-21	15	6	-10	12	-10	21	15	-3	8	34	9	24	21	48	29	100

Fig. 12.6 Values of the correlation coefficients (multiplied by 100) for all pairs of template spectra of single amino acids, calculated in the range from 880 to 1280 cm^{-1} (*right* panel) and from 1300 to 1700 cm^{-1} (*left* panel). The ranges represent extreme (in the sense of mean correlation) windows of 400 cm^{-1} of width (see text and Fig. 12.5). For comparison reasons the ranges of *c* are the same as the ones in Figs. 12.3 and 12.4; however, the boundaries are no longer quantiles.It is often convenient to see how a spectrum of a particular amino acid correlates with those of the other ones. To this end we present Table 12.2, where for each spectrum all correlation coefficients with other spectra are sorted columnwise in decreasing order. This gives us a quick view how far is a given spectrum from others. In Table 12.3 we present statistics that summarize quantitatively the data from Table 12.2

Table 12.2 Values of the correlation coefficients for all pairs of template spectra of single amino acids calculated in the range from 300 cm^{-1} to 1700 cm^{-1}, sorted separately columnwise in decreasing order for each template spectrum shown in the first row

Phe	Tyr	Cys	His	Arg	Gly	Pro	Met	Asn	Ala	Glu	Trp	Ser	Lys	Thr	Asp	Gln	Leu	Val	Ile
Ser	Pro	Met	Gly	Lys	His	Leu	Val	Asp	Ser	Asp	Ile	Ala	Leu	Asp	Thr	Ser	Val	Ile	Val
21	25	23	29	26	29	41	34	45	49	42	32	49	48	57	57	42	60	65	65
Tyr	Gln	Asp	Ser	Ile	Gln	Gln	Ile	Val	Leu	Thr	Asn	Gln	Val	Ile	Asn	Leu	Lys	Leu	Leu
11	16	19	20	20	29	30	33	35	40	34	31	42	45	36	45	37	48	60	48
Trp	Leu	Val	Gln	Ser	Ile	Lys	Cys	Trp	Gln	Ile	Thr	Leu	Ile	Glu	Glu	Ala	Ile	Lys	Thr
11	14	17	19	19	26	26	23	31	37	33	30	29	36	34	42	37	48	45	36
Leu	Ser	Glu	Val	Asn	Ser	Tyr	Asp	Gln	Lys	Val	Asp	Ile	Thr	Lys	Val	Ile	Pro	Asn	Lys
6	13	16	17	18	24	25	23	28	33	25	30	29	33	33	33	32	41	35	36
Ile	Ile	Thr	Leu	Gln	Val	Asp	Glu	Thr	Ile	Leu	Val	Trp	Ala	Trp	Trp	Pro	Ala	Met	Met
5	13	13	16	17	17	19	21	26	33	24	26	26	33	30	30	30	40	34	33
Gln	Phe	Leu	Glu	Val	Asp	Ile	Asn	Ile	Val	Asn	Ser	Gly	Pro	Leu	Gln	Val	Gln	Asp	Glu
5	11	13	14	15	17	18	21	24	29	23	26	24	26	29	29	29	37	33	33
Pro	Cys	Asn	Arg	Thr	Asn	Val	Thr	Glu	Trp	Ser	Gln	Glu	Arg	Val	Ile	Gly	Thr	Gln	Ala
3	11	13	14	14	17	16	20	23	23	22	24	22	26	27	25	29	29	29	33
Ala	His	Tyr	Ile	His	Pro	Gly	Trp	Met	Glu	Trp	Leu	Val	Asp	Gln	Leu	Asp	Ser	Ala	Trp
3	5	11	12	14	16	16	19	21	21	21	23	21	23	26	24	29	29	29	32
Cys	Ala	Gln	Met	Glu	Met	Ser	Leu	Arg	Pro	Met	Ala	Phe	Trp	Asn	Met	Asn	Glu	Thr	Gln
−2	5	10	11	14	15	13	16	18	13	21	23	21	21	26	23	28	24	27	32
Val	Val	Trp	Lys	Gly	Lys	Glu	Lys	Gly	Thr	Ala	Lys	His	Gln	Met	Lys	Thr	Asp	Trp	Ser
−3	3	9	10	13	15	13	15	17	11	21	21	20	20	20	23	26	24	26	29
His	Gly	Ile	Ala	Met	Trp	Ala	Gly	Leu	His	Cys	Glu	Arg	Ser	Arg	Pro	Trp	Trp	Glu	Gly
−3	2	9	10	11	13	13	15	13	10	16	21	19	18	14	19	24	23	25	26
Gly	Arg	Lys	Trp	Asp	Arg	Thr	Gln	Cys	Asp	His	Met	Lys	Met	Ser	Cys	Lys	Met	Ser	Asp
−3	2	5	9	10	13	11	13	13	6	14	19	18	15	13	19	20	16	21	25
Arg	Asn	Pro	Thr	Pro	Thr	Arg	His	Lys	Tyr	Gln	Gly	Tyr	Gly	Cys	Gly	His	His	His	Asn
−4	0	3	8	9	12	9	11	8	5	14	13	13	15	13	17	19	16	17	24
Lys	Met	Ser	Asp	Trp	Glu	His	Arg	Ser	Asn	Arg	Phe	Thr	Glu	Gly	Ser	Arg	Tyr	Gly	Arg
−5	−1	−1	7	8	12	6	11	5	5	14	11	13	12	12	11	17	14	17	20
Glu	Trp	Arg	Pro	Leu	Leu	Phe	Ala	Ala	Phe	Pro	His	Pro	His	Pro	Arg	Tyr	Cys	Cys	Pro
−6	−4	−1	6	8	10	3	3	5	3	13	9	13	10	11	10	16	13	17	18
Asn	Glu	Ala	Tyr	Tyr	Tyr	Cys	Ser	His	Met	Lys	Cys	Asp	Asn	Ala	His	Glu	Asn	Pro	Tyr
−7	−7	−1	5	2	2	3	2	4	3	12	9	11	8	11	7	14	13	16	13
Met	Thr	Phe	Asn	Cys	Ala	Trp	Tyr	Tyr	Cys	Gly	Arg	Asn	Cys	His	Ala	Met	Gly	Arg	His
−8	−9	−2	4	−1	−2	−1	−1	0	−1	12	8	5	5	8	6	13	10	15	12
Asp	Lys	His	Phe	Ala	Phe	Asn	Pro	Pro	Arg	Phe	Pro	Met	Phe	Tyr	Phe	Cys	Arg	Tyr	Cys
−8	−9	−5	−3	−1	−3	−1	−3	−1	−1	−6	−1	2	−5	−9	−8	10	8	3	9
Thr	Asp	Gly	Cys	Phe	Cys	Met	Phe	Phe	Gly	Tyr	Tyr	Cys	Tyr	Phe	Tyr	Phe	Phe	Phe	Phe
−9	−9	−5	−5	−4	−5	−3	−8	−7	−2	−7	−4	−1	−9	−9	−9	5	6	−3	5

intensities, one checks in the corresponding database if they fit to the peaks of a known substance. Some tolerance should be allowed for the positions and intensities to accommodate for the measurement precision. Let us first consider identification by position only. If a strong peak in the template spectrum does not show in the mixture, this is a negative clue, i.e., a component corresponding to the template is probably not part of the mixture. Conversely, if there exists a corresponding peak

Table 12.3 Minimum, mean minus standard deviation, mean, mean plus standard deviation, maximum, 1st tertile (3-quantile), 2nd tertile (second through eighth column) of the correlation coefficients between a given template spectrum (first column) and all other template spectra. The rows are sorted with respect to the mean in increasing order

	Min	Mean -StDev	Mean	Mean +StDev	Max	1-Ter	2-Ter
Phe	−9	−8	0	8	21	−5	3
Tyr	−9	−5	4	14	25	−1	5
Cys	−5	−1	8	16	23	−1	11
His	−5	3	11	19	29	7	12
Arg	−4	4	11	19	26	8	14
Gly	−5	4	13	23	29	12	16
Pro	−3	2	14	25	41	6	16
Met	−8	3	14	25	34	11	19
Asn	−7	3	16	30	45	5	21
Ala	−2	1	17	33	49	5	21
Glu	−7	6	18	30	42	14	21
Trp	−4	8	18	29	32	11	23
Ser	−1	7	20	32	49	13	21
Lys	−9	5	20	35	48	12	23
Thr	−9	4	20	36	57	12	26
Asp	−9	4	21	38	57	11	24
Gln	5	14	24	34	42	17	29
Leu	6	11	26	42	60	14	29
Val	−3	10	27	44	65	17	29

in the mixture this is an indication of the presence of a component in the mixture. Let us consider, for example, 10 strongest peaks in each template. Then, if a peak is present in the mixture spectrum we can assign some number of points to the corresponding amino acid, and in the case of absence—minus some other number of points. Let us note that for only 10 strongest peaks considered the absence of a peak in the mixture spectrum is a stronger sign than the presence and should be properly weighted. In this way we apply scoring to each amino acid and can sort them accordingly. In order to decide if a component is present in the mixture we must assume a certain threshold level for the scores. Obviously, the absence of the strongest peak in the mixture is a stronger negative sign than the absence of the tenth strongest peak. Thus, we can refine the scoring, giving more weight to the first case. Let us note that if a given pair of template spectra has many strong peaks that are close to each other, one may be in doubt which corresponding amino acid is part of the mixture. To find out which pairs of templates can cause such confusion, we determined the number of common positions of 10 strongest peaks for each pair of template spectra. We first defined a series of overlapping windows ($300+4n\ cm^{-1}$; $308+4n\ cm^{-1}$), n = 0, 1, 2, …, 349, covering the whole considered spectral range from 300 cm^{-1} to 1700 cm^{-1}. For each template spectrum we looked for peaks within successive windows and assigned the intensity of the strongest one to them. If there was no peak in the corresponding window, we put zero intensity to it. Thus, we

	His	Phe	Pro	Tyr	Glu	Asn	Gly	Arg	Cys	Asp	Lys	Ser	Met	Ala	Thr	Ile	Trp	Gln	Leu	Val
His	10	1	0	1	0	1	1	0	0	1	0	1	1	2	1	1	0	1	0	2
Phe	1	10	1	2	1	0	1	0	0	0	0	1	0	2	0	1	0	3	1	1
Pro	0	1	10	2	1	0	1	0	0	1	2	1	0	1	3	1	2	3	4	1
Tyr	1	2	2	10	0	1	1	2	2	0	0	1	1	1	0	2	1	2	2	3
Glu	0	1	1	0	10	1	3	1	4	2	0	0	3	2	1	3	0	1	3	2
Asn	1	0	0	1	1	10	1	2	2	3	0	2	2	1	2	2	3	2	1	3
Gly	1	1	1	1	3	1	10	2	0	1	1	2	2	1	1	1	2	3	2	3
Arg	0	0	0	2	1	2	2	10	0	2	1	3	3	0	3	2	2	4	0	3
Cys	0	0	0	2	4	2	0	0	10	3	2	0	3	1	1	2	2	0	4	4
Asp	1	0	1	0	2	3	1	2	3	10	0	1	3	1	3	1	4	2	1	2
Lys	0	0	2	0	0	0	1	1	2	0	10	2	1	4	3	2	4	2	4	3
Ser	1	1	1	1	0	2	2	3	0	1	2	10	1	3	2	2	3	3	1	3
Met	1	0	0	1	3	2	2	3	3	3	1	1	10	2	3	2	2	2	1	3
Ala	2	2	1	1	2	1	1	0	1	1	4	3	2	10	2	3	2	2	5	1
Thr	1	0	3	0	1	2	1	3	1	3	3	2	3	2	10	1	4	3	3	2
Ile	1	1	1	2	3	2	1	2	2	1	2	2	2	3	1	10	2	3	2	5
Trp	0	0	2	1	0	3	2	2	2	4	4	3	2	2	4	2	10	2	4	3
Gln	1	3	3	2	1	2	3	4	0	2	2	3	2	2	3	3	2	10	3	3
Leu	0	1	4	2	3	1	2	0	4	1	4	1	1	5	3	2	4	3	10	4
Val	2	1	1	3	2	3	3	3	4	2	3	3	3	1	2	5	3	3	4	10

Fig. 12.7 Numbers of the common positions of 10 strongest peaks for all pairs of template spectra of single amino acids, calculated in the range from 300 to 1700 cm^{-1}. The colours reflect subdivision of the entries into three essentially equal parts defined by ranges $< 0,1 >$, $(1,2 >$, $(2,10 >$ which implies that the boundaries 1, 2 are 3-quantiles (tertiles)

have a series of windows with intensities that form a simplified spectrum. The width of each window is 8 cm^{-1}; it is related to our experimental resolution. In Fig. 12.7 each entry shows the number of common positions (windows) of two sets of 10 strongest peaks (windows) for a given pair of templates; this number can be viewed as a measure of how close both templates are to each other. The number of peaks considered was chosen somewhat arbitrarily. If one decides that the number of representative peaks in the templates is different, it can be changed accordingly. For comparison, Fig. 12.8 presents the numbers of common positions (windows) for the sets of 19 strongest peaks (windows) for each template. The value of 19 was chosen because this is the number of observed peaks (in the range 300–1700 cm^{-1}) in the spectrum of glycine, which has the simplest structure of all amino acids. Because the windows overlap, we avoid the situation where an original peak is assigned to one window only. Such a situation, if present in two compared spectra, could lead to a failure in identification of the same position of two original spectral peaks that are less than 8 cm^{-1} apart. On the other hand, for overlapping windows a regular

	Gly	Phe	Pro	Tyr	Glu	Cys	Ala	Asn	Thr	Ser	Asp	Met	His	Arg	Lys	Gln	Leu	Trp	Ile	Val
Gly	19	3	3	2	4	0	3	2	2	4	1	2	6	3	3	4	3	3	4	3
Phe	3	19	2	6	2	2	3	0	3	4	3	2	3	3	4	6	2	3	5	2
Pro	3	2	19	3	1	0	3	2	7	3	4	1	2	2	5	5	6	5	3	4
Tyr	2	6	3	19	3	2	1	4	3	4	3	3	6	5	4	3	4	3	3	5
Glu	4	2	1	3	19	7	3	4	2	4	4	6	7	3	1	3	7	3	5	6
Cys	0	2	0	2	7	19	3	4	4	5	5	5	5	4	5	2	6	4	6	9
Ala	3	3	3	1	3	3	19	3	7	7	4	4	5	5	4	5	6	3	7	3
Asn	2	0	2	4	4	4	3	19	2	3	7	8	3	4	5	5	3	7	7	6
Thr	2	3	7	3	2	4	7	2	19	4	5	5	2	6	5	5	4	6	4	4
Ser	4	4	3	4	4	5	7	3	4	19	5	3	5	7	3	3	4	4	4	6
Asp	1	3	4	3	4	5	4	7	5	5	19	5	5	3	3	5	3	8	6	4
Met	2	2	1	3	6	5	4	8	5	3	5	19	4	3	6	6	2	8	5	5
His	6	3	2	6	7	5	5	3	2	5	5	4	19	4	3	3	6	3	2	9
Arg	3	3	2	5	3	4	5	4	6	7	3	3	4	19	5	7	4	3	5	9
Lys	3	4	5	4	1	5	4	5	5	3	3	6	3	5	19	6	5	8	4	7
Gln	4	6	5	3	3	2	5	5	5	3	5	6	3	7	6	19	4	5	7	5
Leu	3	2	6	4	7	6	6	3	4	4	3	2	6	4	5	4	19	6	5	9
Trp	3	3	5	3	3	4	3	7	6	4	8	8	3	3	8	5	6	19	8	5
Ile	4	5	3	3	5	6	7	7	4	4	6	5	2	5	4	7	5	8	19	8
Val	3	2	4	5	6	9	3	6	4	6	4	5	9	9	7	5	9	5	8	19

Fig. 12.8 Numbers of the common positions of 19 strongest peaks for all pairs of template spectra of single amino acids, calculated in the range from 300 to 1700 cm^{-1}. The colours reflect subdivision of the entries into three essentially equal parts defined by ranges $< 0,3 >$, $(3,5 >$, $(5,19 >$ which implies that the boundaries 3, 5 are 3-quantiles (tertiles)

situation is when adjacent windows have the same non-zero intensity coming from the same peak, and a situation is possible when two adjacent common windows are related to only two original peaks of the compared templates (in a non-trivial situation they should be related to three or four peaks). Accordingly, we monitored for such occurrences and reduced the number of common windows appropriately. From Figs. 12.7 and 12.8 one can see that the number of common positions for some pairs of template spectra is relatively high, which means that the positional analysis may not give an unequivocal answer to the question which amino acid from the corresponding pair is present in the mixture.

If the analysis using only positions of peaks is inconclusive, one can support it with the analysis of both positions and intensities. To this end we define a correlated pair of windows as a set of two windows corresponding to nonzero intensities for two compared spectra, the intensities being in the same ratio for each spectrum within a given tolerance. A number of correlated pairs can serve as a measure of similarity between the mixture spectrum and a potential component spectrum. In

	Gly	Phe	Pro	Ala	Tyr	His	Thr	Ser	Trp	Asn	Cys	Asp	Lys	Arg	Gln	Glu	Ile	Met	Leu	Val
Gly	171	1	1	0	1	1	0	1	0	0	0	0	1	0	2	2	1	0	0	0
Phe	1	171	0	1	4	1	1	1	2	0	1	0	2	0	4	1	2	1	0	0
Pro	1	0	171	0	0	0	6	0	4	1	0	3	2	1	6	0	0	0	5	4
Ala	0	1	0	171	0	1	5	7	0	0	0	1	1	1	1	1	6	2	6	1
Tyr	1	4	0	0	171	7	1	1	3	2	0	1	1	2	2	1	1	1	2	5
His	1	1	0	1	7	171	0	2	0	1	2	1	1	1	1	6	0	3	5	7
Thr	0	1	6	5	1	0	171	1	2	0	0	3	3	8	7	0	1	2	3	3
Ser	1	1	0	7	1	2	1	171	1	1	2	4	2	9	1	2	1	3	3	4
Trp	0	2	4	0	3	0	2	1	171	5	1	6	3	0	2	0	7	8	3	2
Asn	0	0	1	0	2	1	0	1	5	171	1	9	5	3	6	1	2	9	0	4
Cys	0	1	0	0	0	2	0	2	1	1	171	4	5	1	0	7	5	2	6	14
Asp	0	0	3	1	1	1	3	4	6	9	4	171	0	1	3	1	1	9	1	3
Lys	1	2	2	1	1	1	3	2	3	5	5	0	171	2	6	0	2	6	4	7
Arg	0	0	1	1	2	1	8	9	0	3	1	1	2	171	10	0	4	2	2	9
Gln	2	4	6	1	2	1	7	1	2	6	0	3	6	10	171	1	5	3	2	0
Glu	2	1	0	1	1	6	0	2	0	1	7	1	0	0	1	171	5	12	10	12
Ile	1	2	0	6	1	0	1	1	7	2	5	1	2	4	5	5	171	4	7	12
Met	0	1	0	2	1	3	2	3	8	9	2	9	6	2	3	12	4	171	0	4
Leu	0	0	5	6	2	5	3	3	3	0	6	1	4	2	2	10	7	0	171	19
Val	0	0	4	1	5	7	3	4	2	4	14	3	7	9	0	12	12	4	19	171

Fig. 12.9 Numbers of correlated pairs of 19 strongest peaks for all pairs of template spectra of single amino acids, calculated in the range from 300 to 1700 cm^{-1} for 50% tolerance for the ratios of intensities. The colours reflect subdivision of the entries into three essentially equal parts defined by ranges $< 0,1 >$, $(1,3 >$, $(3,171 >$ which implies that the boundaries 1, 3 are 3-quantiles (tertiles)

Fig. 12.9 we give the number of correlated pairs for each pair of templates for the sets of 19 peaks and 50% tolerance for the ratio of intensities. This number can serve as a measure of similarity between templates, i.e., the smaller the number, the easier determination which one is present in the mixture. Since the windows overlap and the intensities for adjacent windows often have the same value, as was explained above, we must reduce the number of correlated pairs similarly to the case of positional analysis. For example, considering 19 peaks for each template, one can obtain up to 38 windows with nonzero intensities and comparing a template spectrum with itself can yield $\binom{38}{2}$ correlated pairs. This number should be reduced to $\binom{19}{2} = 171$, as is shown in Fig. 12.9. For each pair of templates we cluster correlated pairs of windows in such a way that two of them with the same intensities and adjacent positions belong to the same cluster. Figure 12.9 reflects the numbers of clusters defined in this manner.

Table 12.4 Qualitative composition of mixtures and their identification numbers

Mixture	Composition
a	Phe Ala
b	His Arg Pro
c	Tyr Asn
d	Gly Thr Ser Gln Leu
e	Trp Glu Gln Ile
f	Ser Thr Arg Ala
g	Met Ala His Gly Leu
h	Glu Leu Ile
i	His Gly Leu
j	Glu Met Lys
k	His Gly Tyr Pro
l	Thr Ala
m	Asn
n	Met Val Leu Ile
o	Phe Gln
p	Tyr Thr Pro Asn Asp
q	Gly Glu Ala
r	Tyr Trp His Arg Ser
s	Trp Asp
t	Tyr Cys Phe His Ser Leu Thr Ile

12.3.3 Component Identification by the Correlation Function

In what follows we present examples of qualitative analysis of mixtures of solid amino acids using as templates the experimentally measured Raman spectra of the components. We assume the linear model of the spectrum of a mixture, which means that it is a linear combination of some number of non-interacting component spectra. We try to find the optimum scaling coefficients of the combination maximizing the correlation function between measured and reconstructed spectra:

$$\sum_{i=1}^{981} \frac{(R_L(w_i) - E(R_L))(M_k(w_i) - E(M_k))}{\|R_L - E(R_L)\|\|M_k - E(M_k)\|}, \tag{12.1}$$

where:

$R_L = \sum_{l \in L} c_l a_l$, $L \subset \{1, 2, 3, \ldots, 20\}$ reconstructed spectrum;

a_l, $l = 1, \ldots, 20$ template spectrum number l (see Table 12.1);

$M_k, k \in \{a, b, c, \ldots, t\}$ measured mixture spectrum identified by Table 12.4;

$w_i = 299 + i, i = 1, 2, 3, \ldots, 981$ wavenumbers (selected optimum range: 300 to 1,280 cm^{-1}, *vide supra*)

$\| \ \|$ Euclidean norm; $E -$ the mean of a vector.

The formula above represents the Pearson correlation coefficient for the reconstructed R_L and measured M_k spectra treated as vectors of intensities for

wavenumbers w_i, and obtained by linear interpolation of the original data points. In the first step of the algorithm we look for a template spectrum most similar to the measured one and accept it as a component spectrum. Then, we try to fit the linear combinations of the spectra of the accepted and the next potential components to the experimentally measured mixture spectrum.

Now, let us assume that we have already accepted some number of component spectra forming a reconstructed spectrum and we want to find out if another template spectrum can be included in it. To this end, we consider a series of two-dimensional optimization problems by looking for the coefficients of the linear combination of the spectrum reconstructed in the previous step treated as a whole, and that of the next potential component. The algorithm proceeds until the stop condition is fulfilled, which means that the difference in the value of the correlation function for the best and worst fit divided by the value for the best fit, or the value of the scaling coefficient for the next best potential component fall below some threshold levels. All optimization tasks were performed under the assumption of non-negativity of components: $c_l \geq 0$ and $\sum_{l \in L} c_l = 1$, which implies that the scaling coefficients cannot be equal to zero simultaneously. Otherwise the denominator in the formula above would be zero. The correlation function is non-concave and, consequently, local maxima may be present. Hence, standard gradient methods for finding the optimum solution may not be suitable. We chose a more general stochastic method for finding the global extremum named Differential Evolution [74], which is a representative of a wider class of genetic algorithms. The method is simple and competitive to other methods from this class, though it cannot guarantee the best solution, due to its heuristic nature. Alternatively, we could have considered a linear combination of all 20 template spectra identified by Table 12.1, which means searching for 20 scaling coefficients at a time. Using such approach, the computing time would be long and the danger of reaching a local maximum would be increased. Therefore it is much safer to consider a series of one- and two-dimensional problems.

The method was tested using twenty mixtures, each containing between 1 and 8 amino acids; the qualitative compositions are presented in Table 12.4. We used solid powders and mixed them in approximately equal volumes. Naturally, this does not mean that all components contribute equally to the mixture spectrum, as different substances are different Raman scatterers. The results of the analysis are presented in Table 12.5. All mixtures were analysed using the same stop condition in the algorithm, assuming that the difference in the value of the correlation function (12.1) for the best and the worst matching template spectrum was less than 1 % of the value for the best match, or, alternatively, the value of the scaling coefficient for the best matching template in the current step is less than 0.03.

The components identified but not present in the mixture are false positives, whereas those not identified but present in the mixtures are false negatives (see Table 12.5). The analysis resulted in 4 false positives and 11 false negatives.

Table 12.5 Identification results by the correlation function of measured mixture spectra denoted by letters; see Table 12.4. All amino acids in the second column were identified correctly using mechanical stop condition of the iterative algorithm (see Sect. 12.3.3). The third column shows the false positive cases, i.e., the components identified but not present in the mixture. The last column shows false negative cases, i.e., the components not identified but present in the mixture

Mixture	Identified correctly	False positives	False negatives
a	Phe		Ala
b	His Arg Pro		
c	Tyr Asn		
d	Gly Thr Ser Gln Leu		
e	Trp Glu Gln		Ile
f	Ser Arg Thr		Ala
g	Met Ala Gly His	Pro	Leu
h	Glu Leu Ile		
i	His Gly	Pro	Leu
j	Glu Met		Lys
k	Gly Tyr His Pro		
l	Thr Ala		
m	Asn		
n	Met Val Leu Ile		
o	Phe	Trp Ser	Gln
p	Tyr Thr Asn Pro Asp		
q	Gly Glu		Ala
r	Tyr Trp Arg His		Ser
s	Trp Asp		
t	Tyr Cys Phe Ser His Thr		Leu Ile

12.3.4 Component Identification by Canberra Distance and Second Order Analysis

In an effort to improve the results of the previous method, we assumed that some problems may come from the subtraction of background from the spectra. The curve enclosing background was drawn manually and for some intervals of wavenumbers we were not sure how to define it exactly. If we assume that this curve varies slowly in comparison with the curve representing the spectrum then it is useful to consider second order derivatives of spectra. Since the spectra are given by sequences of measured data points, the derivatives must be calculated numerically. To this end we used Savitzky- Golay filtering, which relies on fitting a polynomial of some degree to successive local windows of data points. The value of the second derivative of the polynomial in the middle of a window is approximately equal to the second derivative of the function representing the spectrum. This time we try to find the optimum scaling coefficients of the reconstructed spectrum minimizing the following function:

$$\sum_{i=1}^{1401} \frac{2\left|R_L^{(2)}(w_i) - M_k^{(2)}(w_i)\right|}{\left|R_L^{(2)}(w_i)\right| + \left|M_k^{(2)}(w_i)\right|} \tag{12.2}$$

where:

$R_L^{(2)} = \sum_{l \in L} c_l a_l^{(2)}$, $L \subset \{1, 2, 3, \ldots, 20\}$ reconstructed second order derivative of mixture spectrum;

$a_l^{(2)}$, $l = 1, \ldots, 20$ second order derivative of template spectrum number l (see Table 12.1);

$M_k^{(2)}$, $k \in \{a, b, c, \ldots, t\}$ second order derivative of the measured mixture spectrum identified by Table 12.4;

$w_i = 299 + i$, $i = 1, 2, 3, \ldots, 1401$ wavenumbers.

The above function, the distance between objects ($R_L^{(2)}$ and $M_k^{(2)}$ in our case) has been introduced in [75] as Canberra distance. It has been studied theoretically [76], and it is presently applied in genomics as a measure of similarity [77]. Let us observe that if there is a difference in sign between $R_L^{(2)}$ and $M_k^{(2)}$ for some w_i, the corresponding term in (12.2), which can be viewed as local error, attains its maximum value. Obviously this term equals zero if the values $R_L^{(2)}$ and $M_k^{(2)}$ are the same. If they have the same sign the local error approximates maximum value when the difference between them is relatively large. The behaviour of the local errors can be interpreted in such a way that the function (12.2) is more sensitive to the positions and widths of peaks of compared spectra than to the intensities. Moreover, if we multiply $R_L^{(2)}$ and $M_k^{(2)}$ by some constant, the value of (12.2) will not change. This means that the intervals for which $R_L^{(2)}$ and $M_k^{(2)}$ have large values contribute to the function (12.2) similarly as the intervals of small values. We believe that the above features of (12.2) contribute to the quality of the method.

Again, the method was tested qualitatively with the same mixtures as in the case of using function (12.1) (see Table 12.4) and for the same stop condition (see previous Section). The components identified correctly, false positives and false negatives are given in Table 12.6. The analysis yields 5 false positives and only 1 false negative, a much more impressive result than that obtained using the first method.

12.4 Summary and Conclusions

The set consisting of the Raman spectra of twenty proteinogenic amino acids, all recorded under the same experimental conditions, can be used as a basis for the analysis of mixtures. The success (or failure) of such an analysis will depend not only on the number and relative proportions of the components; it is crucially important which particular components are present in the mixture. The results of similarity analysis shown in this work may help in distinguishing between analytically simple and difficult cases. This, in turn, may indicate whether, for a particular case,

Table 12.6 Identification results by the Canberra distance of second order derivatives of measured mixture spectra denoted by letters; see Table 12.4. All amino acids in the second column were identified correctly using mechanical stop condition of the iterative algorithm (see Sect. 12.3.4)

Mixture	Identified correctly	False positives	False negatives
a	Phe Ala		
b	His Arg Pro		
c	Tyr Asn		
d	Gly Thr Ser Gln Leu		
e	Trp Glu Gln Ile		
f	Ser Thr Arg Ala		
g	Met Ala His Gly Leu		
h	Glu Leu Ile		
i	His Gly Leu		
j	Glu Met Lys		
k	His Gly Tyr Pro		
l	Thr Ala		
m	Asn		
n	Met Val Leu Ile	Glu	
o	Phe Gln		
p	Tyr Thr Pro Asn Asp	Gly	
q	Gly Glu Ala		
r	Tyr Trp His Arg Ser	Thr	
s	Trp Asp	Thr Phe	
t	Tyr Cys Phe His Ser Leu Thr		Ile

additional procedures are necessary for reliable determination of the mixture components. As can be deduced from Table 12.6, the initial attempts presented above are quite encouraging. For the four samples consisting of five-component mixtures, in two cases all the components were correctly identified without false positives and negatives. The remaining two cases gave only one false positive, but all the components present were correctly determined. The identification was 100% correct, with zero false positives and negatives for all five cases of three-component mixtures. Finally, for an extremely complicated case of a mixture containing as many as eight amino acids, seven could be identified. The only false negative was isoleucine. Failure of its detection may be due to a simultaneous presence of leucine, which the results of the similarity analysis place very close to isoleucine. Still, a more careful choice of the spectral interval could probably lead to the analytical separation of these two amino acids, as can be inferred from Figs. 12.4 and 12.6. A more general procedure can be envisaged. For example, first, we could define different subregions of wavenumbers for each pair of template spectra, by optimizing a comparison measure. Then, after an initial choice of components using the algorithm described above we could focus on those with small scaling coefficients in the reconstructed spectrum of the mixture. In the next step, we could check which spectra of the remaining, non-identified amino acids are the most similar to the chosen ones in the whole range of wavenumbers, where the analysis was performed initially. Finally, we could substitute successively the spectra of the selected amino

acids with the most similar ones and try to fit the reconstructed spectrum for the corresponding optimal subregions.

Another, still unexplored possibility, is to take into account the high energy region above 2500 cm^{-1}. Figs. 12.1 and 12.2 clearly show significant spectral differences in this range. We plan to analyse this wavenumber interval using similarity analysis techniques described in this work.

One should be rather cautious while trying to interpret the results of statistical similarity analysis in terms of structural/chemical similarities. It should be mentioned, however, that the amino acids which exhibit largest statistical similarities are also strongly chemically similar. This is nicely illustrated, e.g., by the neighbouring positions of valine, leucine and isoleucine in the similarity matrix shown in Fig. 12.3.

Acknowledgements The research was supported by the European Union within European Regional Development Fund, through grant Innovative Economy (POIG.01.01.02-00-008/08).

References

1. Aroca R (2006) Surface-enhanced vibrational spectroscopy. Wiley, Chichester
2. Le Ru EC, Etchegoin PG (2008) Principles of surface-enhanced Raman spectroscopy. Elsevier Science
3. Zhu G, Zhu X, Fan Q, Wan X (2011) Spectrochim Acta A 78:1187
4. Boldyreva E (2007) In: Boeyens JCA, Ogilvie JF (eds) Models, mysteries and magic of molecules. Springer. p 167
5. Tanabe K, Saeki S (1975) Anal Chem 47:118
6. Li J, Hibbert DB, Fuller S, Vaughn G (2006) Chemometrics and intelligent laboratory systems 82:50
7. Culka A, Jehlicka J, Edwards HGM (2010) Spectrochim Acta A 77:978
8. Destrade C, Garrigou-Lagrange C, Forel MT (1971) J Mol Struct 10:203
9. Stenback H (1976) J Raman Spectrosc 5:49
10. Machida K, Kagayama A, Saito Y, Kuroda Y, Uno T (1977) Spectrochim Acta A 33:569
11. Kakihana M, Akiyama M, Nagumo T, Okamoto M (1988) Z Naturforsch A: Phys Sci 43:774
12. Stewart S, Fredericks PM (1999) Spectrochim Acta A 55:1641
13. Kumar S, Rai AK, Singh VB, Rai SB (2005) Spectrochim Acta A 61:2741
14. Murli C, Thomas S, Venkateswaran S, Sharma SM (2005) Physica B 364:233
15. Fukushima K, Onishi T, Shimanouchi T, Mizushima S (1959) Spectrochim Acta 236
16. Wang CH, Storms RD (1971) J Chem Phys 55:3291
17. Machida K, Kagayama A, Saito Y, Uno T (1978) Spectrochim Acta A 34:909
18. Susi H, Byler DM (1980) J Mol Struct 63:1
19. Forss S (1982) J Raman Spectrosc 12:266
20. Migliori A, Maxton PM, Clogston AM, Zirngiebl E, Lowe M (1988) Phys Rev B: Condens Matter 38:13464
21. Kettle SFA, Lugwisha E, Vorderwisch P, Eckert J (1990) Spectrochim Acta A 46:921
22. Gong Y, Wang W-q (2006) Guangpuxue Yu Guangpu Fenxi 26:90
23. Vijayan N, Rajasekaran S, Bhagavannarayana G, Babu RR, Gopalakrishnan R, Palanichamy M, Ramasamy P (2006) Cryst Growth Des 6:2441
24. Gong Y, Zhao J, Chin BY, Wang W (2008) Proc China Assoc Sci Technol 4:187

25. Goncalves RO, Freire PTC, Bordallo HN, Lima JA Jr, Melo FEA, Mendes FJ, Argyriou DN, Lima RJC (2009) J Raman Spectrosc 40:958
26. Belo EA, Lima JA Jr, Freire PTC, Melo FEA, Mendes FJ, Bordallo HN, Polian A (2010) Vib Spectrosc 54:107
27. Alabanza AM, Pozharski E, Aslan K (2012) Cryst Growth Des 12:346
28. Dhanaraj G, Srinivasan MR, Bhat HL (1991) J Raman Spectrosc 22:177
29. Lima RJC, Freire PTC, Sasaki JM, Melo FEA, Mendes FJ (2002) J Raman Spectrosc 33:625
30. Haussuehl S, Karapetyan HA, Sukiasyan RP, Petrosyan AM (2006) Cryst Growth Des 6:2041
31. Aliaga AE, Garrido C, Leyton P, Diaz FG, Gomez-Jeria JS, Aguayo T, Clavijo E, Campos-Vallette MM, Sanchez-Cortes S (2010) Spectrochim Acta A 76:458
32. Faria JLB, Freire PTC, Goncalves RO, Melo FEA, Mendes FJ, Lima RJC, Moreno AJD (2010) Braz J Phys 40:288
33. Casado J, Lopez NJT, Ramirez FJ (1995) J Raman Spectrosc 26:1003
34. Bento ICV, Freire PTC, Melo FEA, Mendes Filho J, Moreno AJD, Joya MR, Pizanic PS (2007) Solid State Commun 141:29
35. Lopez NJT, Hernandez V, Ramirez FJ (1994) Biopolymers 34:1065
36. Madec C, Lauransan J, Garrigou-Lagrange C (1980) Can J Spectrosc 25:47
37. Susi H, Byler DM, Gerasimowicz WV (1983) J Mol Struct 102:63
38. Pawlukojc A, Leciejewicz J, Ramirez-Cuesta AJ, Nowicka-Scheibe J (2005) Spectrochim Acta A 61:2474
39. Mink J, Hajba L, Mihály J, Németh C, Pálmai M, Sandström M (2012) Appl Spectrosc Rev 47:415
40. Dhamelincourt P, Ramirez FJ (1993) Appl Spectrosc 47:446
41. Wang W-N, Wang G, Zhang Y (2011) Chin Phys B 20:123301/1
42. Dhamelincourt P, Ramirez FJ (1991) J Raman Spectrosc 22:577
43. Luz-Lima C, De SGP, Lima JA Jr, Melo FEA, Mendes FJ, Polian A, Freire PTC (2012) Vib Spectrosc 58:181
44. Almeida FM, Freire PTC, Lima RJC, Remédios CMR, Mendes Filho J, Melo FEA (2006) J Raman Spectrosc 37:1296
45. Sabino AS, De SGP, Luz-Lima C, Freire PTC, Melo FEA, Mendes FJ (2009) Solid State Commun 149:1553
46. Krishnan RS, Sankaranarayanan VN, Krishnan K (1973) J Indian Inst Sci 55:66
47. Bougeard D (1983) Ber Bunsenges Phys Chem 87:279
48. Guicheteau J, Argue L, Hyre A, Jacobson M, Christesen SD (2006) Proc SPIE-Int Soc Opt Eng 6218:62180O/1
49. Façanha Filho PF, Freire PTC, Lima KCV, Mendes Filho j, Melo FEA (2008) Braz J Phys 38:131
50. Façanha Filho PF, Freire PTC, Melo FEA, Lemos V, Mendes FJ, Pizani PS, Rossatto DZ (2009) J Raman Spectrosc 40:46
51. Facanha FPF, Jiao X, Freire PTC, Lima JA Jr, dos SAO, Henry PF, Yokaichiya F, Kremner E, Bordallo HN (2011) Phys Chem Chem Phys 13:6576
52. Kumar S (2011) Elixir Online J 4996
53. Kumar S, Rai AK, Rai SB, Rai DK (2010) Indian J Phys 84:563
54. Syamala D, Rajendran V, Natarajan RK, Babu SM (2007) Cryst Growth Des 7:1695
55. Aliaga AE, Osorio-Roman I, Garrido C, Leyton P, Carcamo J, Clavijo E, Gomez-Jeria JS, Diaz FG, Campos-Vallette MM (2009) Vib Spectrosc 50:131
56. Abello L, Genet F, Nigretto JM, Lucazeau G (1989) Surf Sci 215:158
57. Lima JA Jr, Freire PTC, Melo FEA, Lemos V, Mendes FJ, Pizani PS (2008) J Raman Spectrosc 39:1356
58. Pawlukojc A, Leciejewicz J, Tomkinson J, Parker SF (2001) Spectrochim Acta A 57:2513
59. Herlinger AW, Long TV II (1970) J Amer Chem Soc 92:6481
60. Podstawka E, Ozaki Y, Proniewicz LM (2004) Appl Spectrosc 58:570
61. Kapitan J, Baumruk V, Kopecky V Jr, Pohl R, Bour P (2006) J Am Chem Soc 128:13451

62. Carcamo JJ, Aliaga AE, Clavijo E, Garrido C, Gomez-Jeria JS, Campos-Vallette MM (2012) J Raman Spectrosc 43:750
63. Madec C, Lauransan J, Garrigou-Lagrange C (1978) Can J Spectrosc 23:166
64. Machida K, Izumi M, Kagayama A (1979) Spectrochim Acta A 35:1333
65. Madec C, Lauransan J, Garrigou-Lagrange C (1979) C R. Hebd. Seances Acad Sci Ser B 288;69
66. Murli C, Vasanthi R, Sharma SM (2006) Chem Phys 331:77
67. Kolesov BA, Boldyreva EV (2007) J Phys Chem B 111:14387
68. Miura T, Takeuchi H, Harada I (1989) J Raman Spectrosc 20:667
69. Gorelik VS, Zlobina LI, Sardarly RM (1990) Kratk Soobshch Fiz 33
70. Mohan S, Puviarasan N, Bakkialakshmi S (1999) Asian J Chem 11:1137
71. Siddiqui SA, Pandey AK, Dwivedi A, Jain S, Misra N (2010) J Chem Pharm Res 2:835
72. Lima Jr JA, Freire PTC, Lima RJC, Moreno AJD, Mendes Filho J, Melo FEA (2005) J Raman Spectrosc 36:1076
73. da Silva JH, Lima Jr JA, Freire PTC, Lemos V, Mendes Filho J, Melo FEA, Pizani PS, Fischer J, Klemke B, Kemner E, Bordallo HN (2009) J Phys: Condens Matter 21:415404
74. Storn R, Price K (1997) J Global Optim 11:341
75. Lance GN, Williams WT (1966) Comput J 9:60
76. Jurman G, Riccadonna S, Visintainer R, Furlanello C (2009) In: Agrawal S, Burges C, Crammer K (eds) Advances in Ranking, NIPS 09 Workshop. p 22
77. Wu J, Gan M, Zhang W, Jiang R (2011) Int J Biosci Biochem Bioinform 1:102

Chapter 13
Vibrational Spectroscopy: Disease Diagnostics and Beyond

Hugh J. Byrne, Kamila M. Ostrowska, Haq Nawaz, Jennifer Dorney, Aidan D. Meade, Franck Bonnier and Fiona M. Lyng

Abstract This chapter outlines some developments in the applications of vibrational spectroscopy for disease diagnostics and demonstrates how the applications of the spectroscopic techniques can be extended to the analysis and evaluation of disease aetiology and the mechanisms of interaction with and the cellular and subcellular responses to, for example, chemotherapeutic agents and nanoparticles. The primary emphasis is on Raman spectroscopy, although some examples are based on infrared absorption spectroscopy. The studies presented are chosen to illustrate how a range of multivariate analytical techniques can be employed to maximize the potential benefits of the complex spectral information obtained from tissue or cells.

Keywords Vibrational spectroscopy • Raman Spectroscopy • Infrared Spectroscopy • Multivariate statistical analysis • Disease diagnostics • Cellular analysis • Nanotoxicology • Chemotherapeutics

13.1 Introduction

Vibrational Spectroscopy is a subset of spectroscopy which analyses vibrations within a molecule (or material). The vibrations are characteristic of the molecular structure and, in polyatomic molecules, give rise to a spectroscopic "fingerprint". The spectrum of vibrational energies can thus be employed to characterise a molecular structure, or changes to it due to the local environment or external factors (e.g. radiation, chemical agents). Vibrational energies fall within the mid Infrared (IR) region of the electromagnetic spectrum and are commonly probed through IR absorption spectroscopy. Following the discovery of IR radiation by Herschel in 1800 [1], initial applications of Infrared absorption spectroscopy were limited to astronomy and astrophysics [2]. In material sciences, significant advances were made by 1900 when Abney and Festing recorded spectra for 52 compounds, correlating

H. J. Byrne (✉) · K. M. Ostrowska · H. Nawaz · J. Dorney · A. D. Meade
F. Bonnier · F. M. Lyng
Focas Research Institute, Dublin Institute of Technology, Camden Row, Dublin 8 Ireland
e-mail: hugh.byrne@dit.ie

M. Baranska (ed.), *Optical Spectroscopy and Computational Methods in Biology and Medicine,* Challenges and Advances in Computational Chemistry and Physics 14,
DOI 10.1007/978-94-007-7832-0_13, © Springer Science+Business Media Dordrecht 2014

absorption bands with molecular structures [3]. Coblentz helped establish IR spectroscopy as a routine analytical tool, cataloguing the spectra of hundreds of substances, both organic and inorganic [3]. Technological developments post world war II aided considerably in establishing IR spectroscopy as a routine laboratory characterisation technique, but none more so than the development of commercial Fourier Transform IR (FTIR) spectrometers in the 1960s and 1970s [4, 5] and FTIR microscopes in the late 1980s [6].

IR spectroscopy is now a routine technique for materials characterisation and has found numerous applications in forensics [7], environmental science [8] and pharmacology [9]. Raman spectroscopy is a complementary technique with origin in the discovery of the Raman effect in 1928 [10], for which C.V. Raman was awarded the Nobel prize in physics in 1930. In 1998 the Raman Effect was designated an ACS National Historic Chemical Landmark, in recognition of its importance in materials and process analysis. Raman spectroscopy remained largely a curiosity until the advent of the laser in the 1960s, and the revolution in Charged Coupled Detector (CCD) arrays in the 1980s and 1990's added to the benefits of high laser source intensities. In addition, the development of narrow band laser line rejection filters meant that the huge losses in signal from traditional triple monochromator systems could be overcome with the combination of a filter set and a single spectroscopic grating. Furthermore, the significant reductions in acquisition time with multichannel signal detection enabled significant improvements in signal to noise ratios [11]. The combination of technology developments led to a new range of Raman spectroscopic microscopes in the 1990s, establishing Raman spectroscopic microscopy as a relatively inexpensive benchtop laboratory tool to rival conventional infrared spectroscopy.

Similar to IR spectroscopy, Raman spectroscopy entails the coupling of incident radiation with molecular vibrations and the resultant Raman spectrum is characteristic of the compound or material. However, whereas IR spectroscopy involves the absorption of radiation, Raman spectroscopy is a scattering technique, whereby the incident radiation couples with the vibrating polarisation of the molecule and thus generates or annihilates a vibration. The differing underlying mechanisms gives rise to a complementarity of the two techniques. For a vibration to be active in IR spectroscopy, a change in dipole is required, whereas to be Raman active, a change in polarisability is required. Vibrations of asymmetric, polar bonds thus tend to be strong in IR spectra, whereas Raman is particularly suitable as a probe of symmetric, nonpolar groups. Notably, OH vibrations of water are very strong in IR spectra, whereas they are extremely weak in Raman spectra, rendering Raman a potentially more suitable technique for bioapplications.

A further implication of the differing physical origins of the techniques is that whereas IR monitors the absorption of IR radiation, Raman scattering can be employed in the UV, visible or near IR regions of the spectrum. Raman scattering thus offers intrinsically higher spatial resolution for mapping or profiling, the diffraction limit being determined by the wavelength (<1 μm for Raman, ~5–10 μm for IR). For many applications, however, near IR is favoured as a source for Raman analysis, to minimise interference from scattering, fluorescence, or photodegradation of the sample [12].

FTIR has become an accepted tool in biophysics for analysis of the structure and interactions of proteins [13], lipids [14], carbohydrates [15] and nucleic acids

[16]. Applications to tissue samples for (cancer) diagnostic applications were first reported in the early 1990s [17], and since this time a range of pathologies has been investigated [18–20]. The application of Raman spectroscopy to biomolecules and even tissues was first demonstrated as early as the 1960s [21–23] and by the mid 1970s biomedical applications were explored [24]. Whole cell and tissue studies have been carried out on a range of pathologies [25–29] and *in vivo* studies [30–32] have demonstrated the prospective for diagnostic applications. The potential of vibrational spectroscopy in conjunction with multivariate analysis techniques as a diagnostic tool has thus been well demonstrated and the concept of Spectral Cytopathology has been coined [33].

This chapter is based on the experience of the research activities within the Focas Research Institute, Dublin Institute of Technology, and outlines efforts and progress in the applications of both Raman and infrared spectroscopy for the routine screening of disease as well as the understanding disease origin. It further describes studies demonstrating the potential of the techniques for *in vitro* diagnostics and biochemical analyses, for example of the efficacy of chemotherapeutic agents and the interaction of nanoparticles in cells.

13.2 Vibrational Spectroscopy for Cervical Cancer Pathology [34]

Cervical cancer is the second most common cancer among women worldwide and generally is more common in developing countries. However, substantial decreases in cervical cancer mortality have been observed in Western countries and some developing countries with well developed screening programmes [35]. A Papanicolaou test (Pap smear test) is used to screen for Cervical Intraepithelial Neoplasia (CIN) and cervical cancer in the general female population. Such cervical cytology can have a high specificity of 95–98 % but a sensitivity of lower than 50 % [36]. Other methods such as automated cytology and human papilloma virus (HPV) testing [37–39] have been introduced to reduce the false negative rates. An abnormal Pap smear is followed by colposcopic examination, biopsy and histological confirmation of the clinical diagnosis. This involves the visual examination of histological sections. The grading characteristics are quite subjective and pre-malignancy may not be visually perceptible at all.

Optical technologies such as fluorescence spectroscopy [40], polarised light scattering spectroscopy [41], optical coherence tomography [42, 43] and confocal reflectance microscopy [44] have emerged in recent years as promising tools for diagnosis of cervical and other cancers. The potential of vibrational spectroscopy for cervical cancer diagnosis has also been recognised. A number of studies have shown that Fourier Transform Infra Red (FTIR) spectroscopy [45–47] or Raman spectroscopy [32, 48] can be used for detecting spectral changes in malignant and pre-malignant cells. Raman spectroscopy is, however, more suited to *in vivo* studies because of minimal interference from water. Coupled with adapted multivariate

analysis techniques, Raman spectroscopic analysis of cervical biopsies and cytological samples can provide high specificity and sensitivity for the detection of cervical cancer and early pre-malignant states. The following provides an illustration of such applications combining Raman analysis and Principal Component Analysis/Linear Discriminate Analysis [34].

Formalin fixed paraffin preserved (FFPP) cervical tissue samples were obtained from the National Maternity Hospital, Holles St., Dublin. Two parallel 10 μm FFPP sections were cut from each block using a microtome, mounted on glass slides and dried. Samples were dewaxed in-house prior to investigation by immersion in baths of Xylene (BDH), Ethanol Absolut (Merck) and Industrial Methylated Spirits 95% (Lennox) and air-dried. Cervical tissue was dewaxed using the same procedure outlined above, but xylene was substituted with hexane, due to the improved level of dewaxing [49]. One section from each sample (the reference section) was stained with Haematoxylin and Eosin and the other kept unstained for spectroscopic examination.

FFPP cervical tissue sections were characterised by consultant pathologist Dr. Kelehan P, National Maternity Hospital, Holles St., Dublin, and the samples consisted of 20 normal and 20 invasive carcinoma sections from 40 individuals. Of the 20 carcinoma samples, 10 samples were identified as having various grades of CIN, which were also marked for examination.

An Instruments S. A. Labram Raman spectroscopic confocal microscope was used, with an Argon Ion laser operating at an excitation wavelength of 514.5 nm. A x50 objective lens was used. The laser power at the sample was measured and found to be 6.5 ± 0.05 mW, focused to a spot size of ≈ 2 μm at the tissue surface. Raman spectra of the various tissues were accumulated for 150 s. A total of 10 spectra were recorded from different spots on each sample.

Selected spectra were baseline corrected, dark current subtracted and were smoothed using a 10 point moving average before being normalised to the spectral maximum, from 0 to 1. Basic matrix manipulations and data reduction was carried out in Microsoft Excel Professional 2003 (v. 11.0), before being exported into Minitab to perform Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). Minitab Release 14.1 Statistical Software Analysis Programme was used to produce PCA scores and LDA plots, as well as to carry out leave-one-out cross validation.

Figure 13.1 shows the layer structure of normal cervical tissue in an unstained FFPP tissue section together with the Raman spectra recorded from basal cells, epithelial cells and connective tissue. The spectra of the three different regions do have a degree of similarity, but notably there are many differentiating spectral features. The spectra of basal cells show strong bands at 724, 779 and 1,578 cm^{-1} which are characteristic of nucleic acids. The morphology of basal cells consists of a single line of tightly packed cells, with large nuclei in relation to the compacted surrounding cytoplasm. In addition, these cells are constantly dividing, providing cells to the parabasal layer. For both of these reasons, a high concentration of DNA would be expected in the basal cells. Spectra of epithelial cells have characteristic glycogen bands at 482, 849, 938, 1,082 and 1,336 cm^{-1}. Collagen contributions can be clearly seen in the spectra of connective tissue at 850, 940 and 1,245 cm^{-1}.

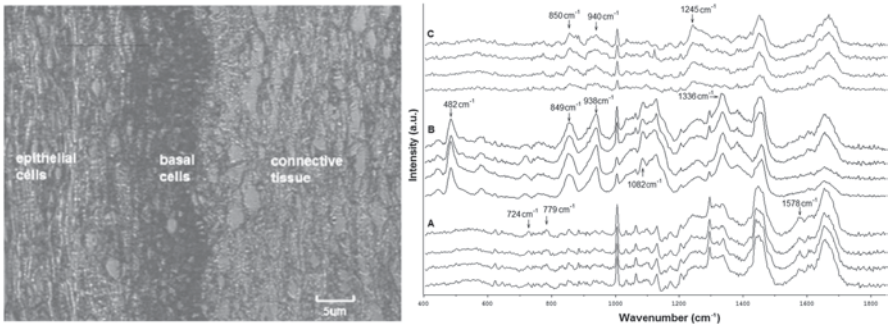


Fig. 13.1 **a** Photomicrograph of unstained tissue section, with different cell types identified. **b** Raman spectra recorded from basal cells (*A*), epithelial cells (*B*), and connective tissue (*C*) in cervical tissue sections. The main spectral features associated with each cell type are highlighted. (Reproduced from [34])

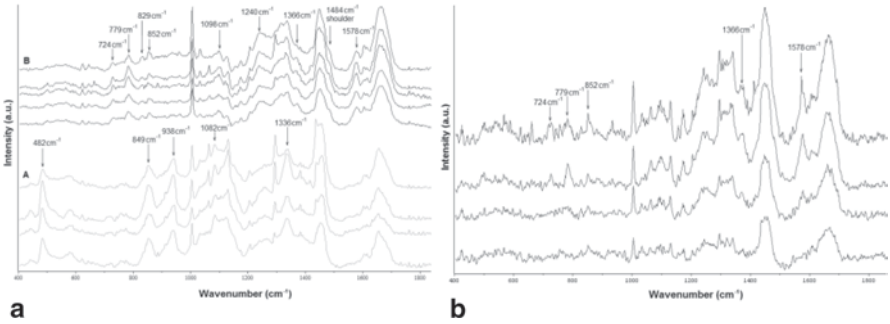


Fig. 13.2 Raman spectra of **a** normal cervical epithelial cells (*A*) and invasive carcinoma cells (*B*) **and (b** Raman spectra of CIN tissue. Assignments of the main Raman vibrational modes are detailed in Table 13.1. (Reproduced from [34])

Figure 13.2a compares the Raman spectra collected from normal epithelial cells and invasive carcinoma from a selection of different patients. Glycogen contributions are clearly visible in the spectra from the normal epithelial tissue. The most obvious bands arise at 482, 849 and 938 cm^{-1} and are due to glycogen skeletal deformation, CCH aromatic deformation and CCH deformation respectively. However, there are also other glycogen contributions not as apparent, including a CC stretching band at 1,082 cm^{-1} and CH_3CH_2 wagging at 1,336 cm^{-1}. These glycogen bands (482, 849 and 938 cm^{-1}) are absent in the spectra from invasive carcinoma, and a reduction in the intensity of the CC stretching mode (1,082 cm^{-1}) is also apparent. Glycogen, a polysaccharide, is the principal storage form of glucose. Cervical cells are unusual among other epithelial cells in that they accumulate large amounts of glycogen during the maturation process. Glycogen is known to be linked with cellular maturation and disappears with the loss of differentiation during neoplasia. The spectra of invasive carcinoma also show characteristic nucleic

Table 13.1 Peak position and assignments of main Raman vibrational modes

Peak position (cm^{-1})	Assignment	Peak reference (Fig. 13.4)
622	C–C twisting (Proteins)	1
642	C–C twisting mode of Tyr and Phe	2
670	C, T,G (DNA/RNA)	3
720	C–N stretching in A and lipids	4
724	CH_2 deformation	
746	CH_2 rocking	
750	Symmetric ring breathing (protein)	5
782	U, T, C (ring breathing modes in the DNA/RNA)	6
827	PO_2 stretching in DNA, Tyr	7
832	CCH deformation aliphatic	
854	Ring breathing in Tyr and Pro (proteins)	8
873	CC stretch	
922	C–C stretching	
935	C–C stretching mode of Pro and Val	9
1004	CC aromatic ring breathing	10
1031	C–H bending mode in Phenylalanine	11
1034	C–N stretching in proteins	
1065	C–N stretch	
1093	Symmetric PO_2 stretching of the DNA backbone; lipids	12
1096	C–C chain stretching	
1098	CC stretch	
1102	CC stretch	
1124	CC skeletal stretch *trans*	
1128	C–N stretching in proteins; C–O stretching in carbohydrates	13
1155	C–C and C–N stretching of proteins	14
1175	C–H in plane bending mode of Tyr and Phe; C, G	15
1208	C–C_6H_5 stretching mode in Trp, Phe;	16
1214	CC stretch backbone carbon phenyl ring	
1220–1280	Amide III; A,C,T ring breathing modes of the DNA/RNA	17
1236	CN stretch, NH bending Amide III band	
1240	CN stretch, NH bending Amide III band	
1311	CH_3/CH_2 twisting mode of collagen and lipid	18
1314	CH deformation	
1337	CH_2 deformation	
1335	CH_2 deformation	
1340	G (DNA/RNA), CH deformation in proteins and carbohydrates; Trp	19
1366	CH_2 bending	
1440	CH_2 scissoring	
1450	CH (CH_2) bending mode in proteins and lipids	20
1484	CH_2 deformation	
1548	NH deformation; CN stretch Amide II band	
1578	C=C olefinic stretch	
1583	A,G (DNA/RNA); C=C bending mode of Phe	21

Table 13.1 (continued)

Peak position (cm^{-1})	Assignment	Peak reference (Fig. 13.4)
1585	C=C stretching	
1550–1700	H_2O bending mode	22
1602	CO stretching	
1618	C=C Phe, Tyr and Trp	23
1660–1665	C=O stretch Amide I a-helix	24
2930	CH_2 stretching (2930 cm^{-1})	
2932	CH3 symmetric stretch (2932 cm^{-1})	

acid bands. These include prominent bands at 724, 779 and 1,578 cm^{-1}, but also at 829 cm^{-1}, 852 cm^{-1}, 1,002 cm^{-1}, 1,098 cm^{-1}, 1,240 cm^{-1}. Distinct bands are also seen at 1,366 cm^{-1}, a shoulder at 1,484 cm^{-1} and a band at 1,578 cm^{-1}. An increase in the intensity of the Amide I band (1,655 cm^{-1}) was observed in the spectra of carcinoma samples compared to the normal tissue samples. The increased nucleic acid and protein bands are a result of the increased proliferation of these tumour cells.

To investigate whether pre-malignant changes can be elucidated using Raman spectroscopy, 10 areas of CIN from 10 different patients were marked by the pathologist and a selection of the resulting Raman spectra are shown in Fig. 13.2b. A number of the spectral features identified in the invasive carcinoma samples were also observed in the CIN samples such as the nucleic acid bands at 724, 779, 852, 1,366 and 1,578 cm^{-1}. This indicates that biochemical changes associated with CIN can be identified using Raman spectroscopy.

Principal component analysis (PCA) is a method of multivariate analysis widely used with datasets of multiple dimensions. PCA is a chemometric technique that can resolve a complete spectral data set into a few principal components and can thus identify and isolate important trends within the data set [50]. It allows the reduction of the number of variables in a multidimensional dataset, although it retains most of the variation within the dataset. The order of the PCs denotes their importance to the dataset. PC_1 describes the highest amount of variation, PC_2 the second greatest and so on. Therefore, var (PC_1)≥var (PC_2)≥var (PC_p), where var (PC_i) represents the variance of PC_i in the considered data set. Generally, the 3 first components represent more than 90 % of the variance. This statistical method was preferred for this study to highlight the variability existing in the spectral data set recorded during the different experiments. Another advantage of this method is the observation of loadings which represent the variance for each variable (wavenumber) for a given PC. Analysing the loadings of a PC can give information about the source of the variability inside a data set, derived from variations in the molecular components contributing to the spectra.

Principal component analysis was used to reduce the number of parameters needed to represent the variance in the spectral data set. The principal components were then used to generate a linear discriminant model. LDA applies a linear discriminant function that maximises the variance in the data between groups and minimises the variance between members of the same group [51]. All three tissue classes were

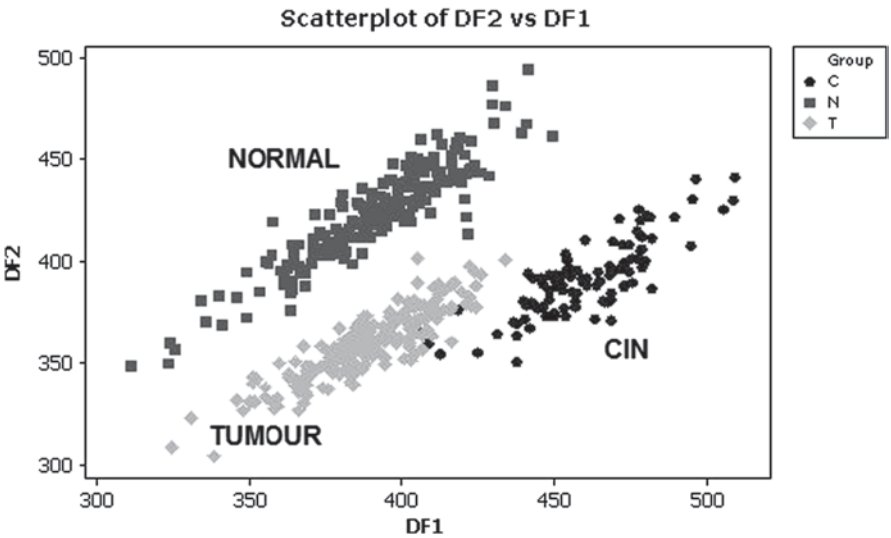


Fig. 13.3 Linear discriminant analysis of the principal components of the first derivative, normalised, 10 point averaged spectra, over the entire spectral range, *C* CIN, *N* normal and *T* invasive carcinoma. (Reproduced from [34])

successfully discriminated as shown in Fig. 13.3. The classification model was tested using a leave one out cross validation in which all but one spectrum was used to build the model. This model was then used to predict the remaining spectrum. This was repeated for all 498 spectra and the results are shown in Table 13.2. Of 498 tissue spectra, 492 were correctly classified as normal, invasive carcinoma or CIN. The cross validation mis-classified 6 spectra, 2 of which were normal samples assigned as invasive carcinoma. The other 4 were either invasive carcinoma or CIN mis-classified as either CIN or invasive carcinoma respectively. Importantly, no abnormal samples were classified as normal. Based on the cross validation results, sensitivity and specificity values were calculated as 99.5 and 100 % respectively for normal tissue, 99 and 99.2 % respectively for CIN and 98.5 and 99 % respectively for invasive carcinoma. It is possible that these values may be slightly over optimistic because of the LDA method used with leave one out cross validation. However, similar methods have been used previously to classify bladder and prostate cancer [52], breast cancer [53] and basal cell carcinoma [54]. Ideally, the use of an unknown test set would provide a more robust analysis of the sensitivity and specificity of the technique and this is planned for future work.

The results show the ability of Raman spectroscopy to classify cervical cancer and pre-cancer with high sensitivity and specificity. These classifications are based on biochemical changes known to accompany cervical cancer such as loss of differentiation and increased proliferation. The study presented details of the application to cervical biopsies, but they can equally be applied to cytological samples. Studies in a range of different pathologies have emerged and technological development is progressing beyond the stage of "proof of concept" [55, 56]. Technologically, *ex-*

Table 13.2 Summary of cross-validation classification results and sensitivity and specificity values

	Diagnosis (Histopathology)		
Predicted (Raman)	Normal	CIN	Invasive carcinoma
Normal	198	0	0
CIN	0	99	3
Invasive carcinoma	2	1	195
Total	200	100	198
Number correct	198	99	195
Proportion	0.990	0.990	0.985
Sensitivity	99.5%	99.0%	98.5%
Specificity	100%	99.2%	99.0%

vivo screening of histological or cytological samples is less challenging, but as the technologies are optically based, the emergence of *in-vivo* probes is inevitable [57]. The development of such an automated technique measuring biochemical changes with improved diagnostic capability could allow faster, more effective patient management and inevitably would increase survival rates. Raman spectroscopy therefore offers enormous potential as a 'next generation' technology to assist pathologists and cytologists with cervical cancer, and other pathological screening and diagnosis.

In the development of the technologies towards clinical and other applications, it is important to ensure that the preprocessing of the spectral data is appropriate and in this context it is important to understand the physical and instrumental origin of any measurement artefacts. Significant progress has been made in FTIR spectroscopy to identify and correct for so-called dispersion artefacts [58–60]. In Raman microspectroscopy, it has been demonstrated that measurement in immersion significantly reduces the spectral background [61] and photodamage [62], minimizing the need for spectral preprocessing.

Notably, the analytical capabilities of the spectral techniques rely heavily on multivariate analytical techniques to classify the complex spectral signatures, and changes to them, associated with disease. A broad range of data analytical techniques are available including Genetic Algorithms, Support Vector Machines, Neural Networks etc. To date, no direct comparison of the efficacy or suitability of such techniques has been made and such a comparison on experimental data may not be conclusive, as the true result is not known. Ultimately, the field would benefit from an exhaustive comparison of the range of data preprocessing and analysis techniques on appropriate simulated datasets [59].

13.3 Exploring the Aetiology of Disease—HPV Infection [63, 64]

The probability of developing cervical dysplasia is much higher for women infected with the human papillomavirus (HPV) than those free of infection [65]. Infection with oncogenic or high-risk HPV (HPV-16, HPV-18) is the main risk factor for

cervical cancer [66], 99.7% of invasive cervical neoplasia being associated with HPV infection [67]. Identifying the presence of HPV is a very important factor in the prevention of cervical cancer and thus, HPV testing has been added to the range of clinical options for cervical cancer screening. HPV testing is based on detecting the DNA of the virus. It has been found that the DNA of high risk strains of HPV is assimilated within the cellular DNA of the host [68]. A drawback to HPV testing is that it is more expensive and time-consuming than other screening tests, and it requires a sophisticated laboratory infrastructure.

Raman spectroscopy has also been used to study cell lines derived from the cervix, revealing spectral variations, mostly in peaks originating from DNA and proteins, in cell lines expressing the E7 gene of HPV-16 compared to cells not affected by HPV [69]. Furthermore, it has been demonstrated that, in low-grade cytology infected with high-risk oncogenic HPV16 or HPV18, it was possible to differentiate samples from women based on whether they were aged 20–29 vs. 30–39 years using attenuated total reflection infrared spectroscopy [70].

To further demonstrate the potential of vibrational spectroscopy for clinical screening, four cervical cancer cell lines were investigated using both Raman and FTIR spectroscopy; HPV negative C33A, HPV-18 positive HeLa with 20–50 integrated HPV copies per cell, HPV-16 positive SiHa with 1–2 integrated HPV strands per cell and HPV-16 positive CaSki containing 60–600 integrated HPV copies per cell. The main aim of the experiment was to investigate biochemical changes in cells caused by high risk HPV strands (HPV-16 and HPV-18) and to examine whether it is possible to differentiate between the cells with no HPV copies and cells infected by low number, medium number and high number of integrated HPV copies, using vibrational spectroscopic techniques. Both common forms of vibrational spectroscopy, Raman and Fourier Transform Infrared absorption, were used to gain complementary information [63]. The Raman spectroscopic analysis is discussed here.

A Horiba Jobin Yvon LabRam 800HR was employed for the measurement, operating at a wavelength of 532 nm and utilising a x100 objective. After incubation, cells were centrifuged to form pellets and washed thoroughly before measurement. The mean of 30 Raman spectra recorded from two pellets (approximately 15 spectra per pellet) for C33A, SiHa, HeLa and CaSki cell lines were derived from the raw spectra via the pre-processing procedures and are presented in Fig. 13.4. Assignments of the main peaks (numbered in Fig. 13.4) are shown in Table 13.1. In the mean Raman spectra, almost no significant differences between HPV negative C33A and SiHa cells, containing 1–2 HPV copies per cells (low HPV concentration) are discernible. Only the peak at 1,583 cm^{-1} assigned to DNA/RNA and Phenylalanine vibrations was found to vary between C33A and SiHa cells. For cells containing a higher number of integrated HPV copies per cell, HeLa cells (medium HPV concentration) and CaSki cells (high HPV concentration), the observed differences are more prominent for protein, nucleic acid and lipid vibrations. CaSki cells exhibited a significant increase in Amide I (proteins) intensity at 1,660 cm^{-1} and a decrease in lipid levels at 1,093, 1,311 and 1,450 cm^{-1} compared to C33A, SiHa and HeLa. HPV presence in a cell appears to change protein expression, which may result in the protein variations observed in the Raman spectra, primarily for

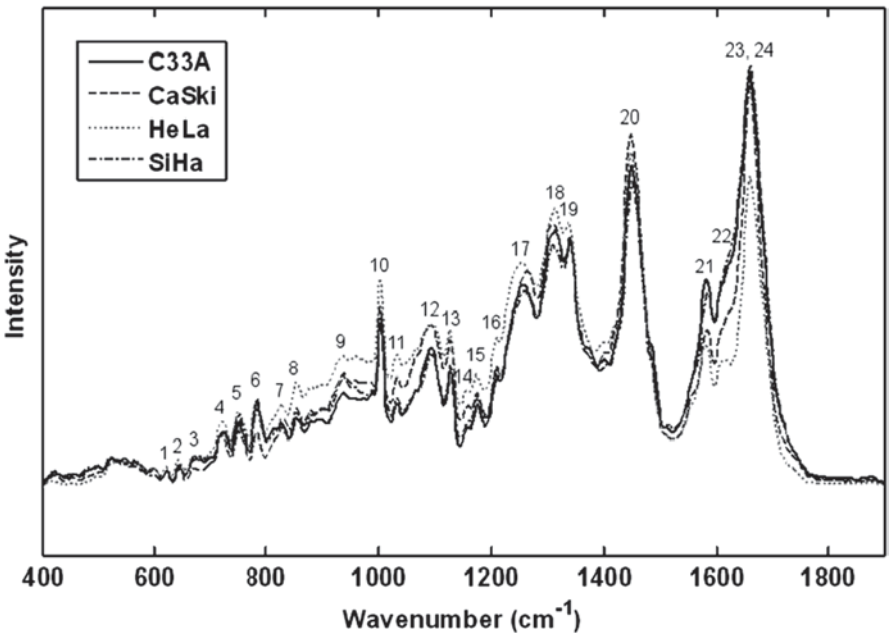


Fig. 13.4 Mean Raman spectra of cervical cancer cell lines. (Reproduced from [63])

the Amide I region. Further differences in protein vibrations between the cells affected by non- and low HPV number compared to medium and high HPV number were observed in the Amide III region (1,254 cm^{-1}), and protein constituent amino acids Tryptophan (Trp), Proline (Pro), Tyrosine (Tyr) and Phenylalanine (Phe) vibrations at 854, 935, 1,003, 1,583 and 1,618 cm^{-1}. Variations in vibrations associated with nucleic acid constituents: PO_2 group, A, C, T, U at 1,093, 720, 782, 1,254, 1,583 cm^{-1} respectively were also observed. The differences in protein and nucleic acid composition support the molecular data which demonstrates intergration of HPV DNA at chromosome fragile sites and the subsequent downstream effect of HPV E6 and E7 oncoprotein expression on host cell proteins such as p53 [71].

The principal component (PC) scatterplot for Raman spectroscopic results presented in Fig. 13.5a shows a distinctive separation between the cell lines, suggesting a HPV concentration dependence factor. The C33A (non HPV) and SiHa (low HPV concentration, LC) cell lines are grouped together, confirming observed similarities between these two types of cells based on mean spectra analysis. The HeLa (medium HPV concentration, MC) and CaSki (high HPV concentration, HC) cell lines are clearly separated from each other and from the C33A-SiHa group. Although the 3D PCA scatterplot demonstrates the HPV concentration dependent separation, the cell lines are not grouped according to increasing (or decreasing) number of HPV copies. CaSki (HC) cells lie between HeLa (MC) and SiHa (LC). This suggests the influence on the data point distribution of another factor—HPV strand type. HeLa are cells infected by HPV-18, while CaSki and SiHa are infected by

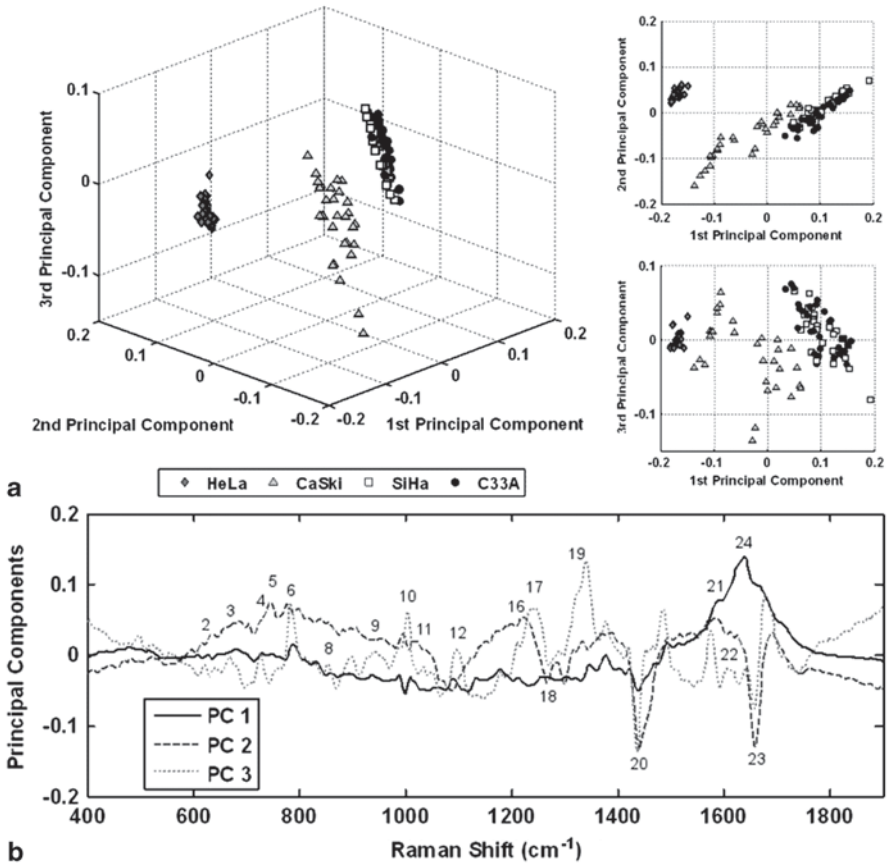


Fig. 13.5 Principal component analysis for Raman spectroscopy results for cervical cancer cell lines. **a** Three dimensional PCA scatterplot. Graphs show separation of the cervical cancer cell lines suggesting HPV concentration dependence. **b** Principal component loadings. Peak assignments correspond to those listed in Table 13.1. (Reproduced from [63])

HPV-16. Principal component loadings are presented in Fig. 13.5b. Peak numbers correspond to assignments presented in Table 13.1. The percent of variance explained by PC1, PC2 and PC3 is 71.9, 10.8 and 6.2%, respectively. The first principal component is dominated by the contribution of the H-O-H bending of water, a broad peak in the range of 1,550–1,700 cm^{-1} with maximum at ~1,644 cm^{-1} [72]. Water is not, however, the only potential contribution within this range, as Amide I vibrations (β-sheet) exhibit a maximum at 1,640 cm^{-1}. Nucleic acid contributions to PC1 may also be present at 782, 1,093 and 1,583 cm^{-1}, although it is very difficult to judge what is the percentage of contribution of each of these components. The second principal component shows predominantly protein and lipid contributions. Lipids are exhibited at 1,095, 1,310 and 1,450 cm^{-1}, while proteins are at 1,660 cm^{-1}-Amide I (α-helix), 1,270 cm^{-1}-Amide III (α-helix), 1,450 cm^{-1}, C=C

Phe at 1,582 cm^{-1} and Trp at 750 cm^{-1}. The third principal component contains a lot of biological information with mainly nucleic acid and protein contributions but only accounts for 6.2% of variance. There are distinctive peaks corresponding to nucleic acids and protein vibrations. DNA and RNA contributions are observed at 670, 720, 782, 1,092, 1,250, 1,340 and 1,583 cm^{-1}. PC3 shows a strong influence of variation arising from protein vibrations, Phe (642, 1,002, 1,031, 1,583 and 1,618 cm^{-1}), Trp (1,208, 1,340 and 1,618 cm^{-1}), Tyr (642, 854 and 1,618 cm^{-1}), Pro (854 cm^{-1}, 935 cm^{-1}), Val (935 cm^{-1}), Amide I at 1,680 cm^{-1} (disordered) and Amide III at 1,260 cm^{-1} (disordered). The source of variations shown in the loadings is in very good agreement with those observed in the analysis of the mean spectra for the investigated cell lines. The main peaks in the loadings arise from protein, nucleic acid and lipid component vibrations. Each principal component is dominated by different cellular components; water, and possibly nucleic acids and proteins in PC1, lipids and proteins in PC2 and nucleic acids and proteins in PC3. Amide I and Amide III contributions are present in all principal component loadings indicative of differences in structural conformation of the protein β-sheet in PC1, α-helix in PC2 and disordered conformation in PC3. As mentioned previously, variations in protein, nucleic acid and lipid levels arise from the fact that the viral DNA interacts with the cellular DNA, causing reprogramming of protein expression in the cell. It was noted that the multivariate analysis could be sustantially influenced by water contributions in the region of 1,640 cm^{-1}. Repeated over the spectral range 400–1,550 cm^{-1}, the PCA gave a similar separation of the cellular types and yielded similar principal components.

The applications of spectroscopic analysis, coupled with multivariate statistic techniques can therefore be extended beyond diagnosis based on classification to biochemical analysis of the underlying physiology based on changes to the molecualr composition on a cellular and subcellular level. This new technology for cervical screening not only can distinguish cervical cells based on biochemical content but also can elucidate the effect of HPV on the cells. As a further demonstration of the potential of the techniques, FTIR spectroscopy was employed to further explore the known manifestations of HPV infection in cervical cell lines [64].

Differentially expressed proteins in cancer have potential utility as biomarkers. As the cell cycle is often disrupted in a cancerous cell, proteins associated with it are often candidate biomarkers. Putative biomarkers of cervical cancer that are currently under study include proteins such as CDC 6 (DNA licensing protein), minichromosome maintenance proteins (MCM 2, MCM 5), p53 or $p16^{INK4A}$ [73]. These biomarkers have been used to detect the presence of abnormal cells, based upon immunocytochemical methods. $p16^{INK4A}$ regulates the levels of active cyclin D/CDK in the cell, providing a feedback mechanism that regulates the levels of MCM (minichromosome maintenance proteins), PCNA (proliferating-cell nuclear antigen) and cyclin E. Overexpression of $p16^{INK4A}$, which is considered a marker of elevated E7 expression, can be detected in some CIN1 lesions, as well as in CIN2 and CIN3 lesions that show evidence of integration [74, 75]. Integration of the HPV genome into the host cell chromosome is a critical event in the development of most cervical cancers [76]. Overexpression of $p16^{INK4A}$ has been demonstrated in

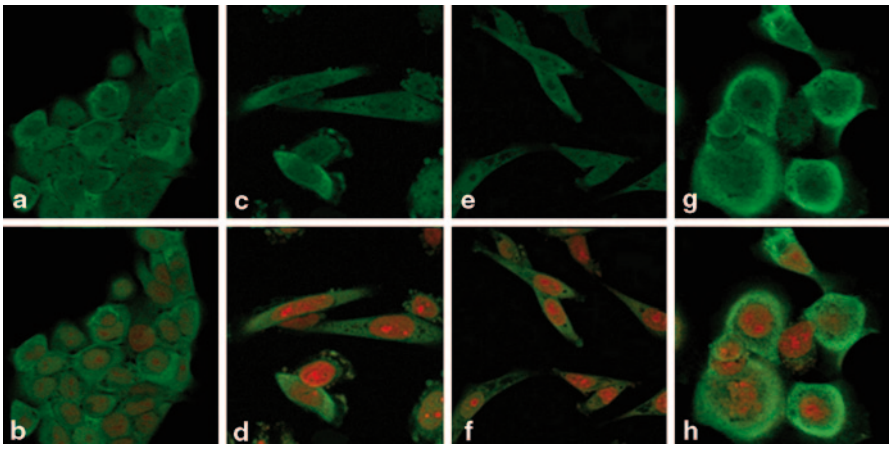


Fig. 13.6 Confocal microscopy images of cervical cell lines: **a**–**b** C33A cells. **c**–**d** SiHa cells. **e**–**f** HeLa cells. **g**–**h** CaSki cells. Images **a**, **c**, **e** and **g** show FITC fluorescence staining (related to p16INK4A expression), while images **b**, **d**, **f** and **h** present FITC and PI staining. (Reproduced from [64])

cervical cancers as a result of functional inactivation of pRb by the HPV E7 protein [77]. This overexpression highlights the potential of $p16^{INK4A}$ as a marker for CIN and cervical cancer. HPV positivity and $p16^{INK4A}$ positivity have shown a correlation, even though $p16^{INK4A}$ expression was also seen in a limited number of HPV negative biopsy samples [78]. Also, a correlation between $p16^{INK4A}$ expression and cervical lesion grade and HR-HPV positivity has been documented [79]. $p16^{INK4A}$ has been proven to be the most reliable marker of cervical dysplasia and was found to mark all grades of squamous and glandular lesions of the cervix. The use of $p16^{INK4A}$ immunocytochemical analysis as a complement to conventional screening programmes could potentially aid in the reduction of false positive and false negative results [80].

In the first part of this study, expression of $p16^{INK4A}$ was analysed in cervical cancer cell lines using immunocytochemical staining and both confocal fluorescence microscopy and flow cytometry. [64] Fluorescence images of cells were collected using a confocal fluorescence microscopy system. Representative images of cells are presented in Fig. 13.6, where a, c, e and g show only green fluorescence staining (FITC), while b, d, f and h show the overlay of green (FITC) and red (PI) fluorescence signals. The FITC fluorescence intensity is associated with the $p16^{INK4A}$ expression level, while the PI dye was used to visualise the position of nuclei within the cells.

A degree of correlation between the number of HPV copies per cell line and the fluorescence intensity related to $p16^{INK4A}$ expression was observed for the flow cytometry data (Fig. 13.7). The highest level of $p16^{INK4A}$ expression (intensity of staining) was observed for CaSki cells, followed by HeLa, SiHa and the lowest for C33A. The relationship between $p16^{INK4A}$ expression and HPV copy number, illustrated by Fig. 13.7, is supported by a study conducted by Klaes et al. (2001), wherein a corre-

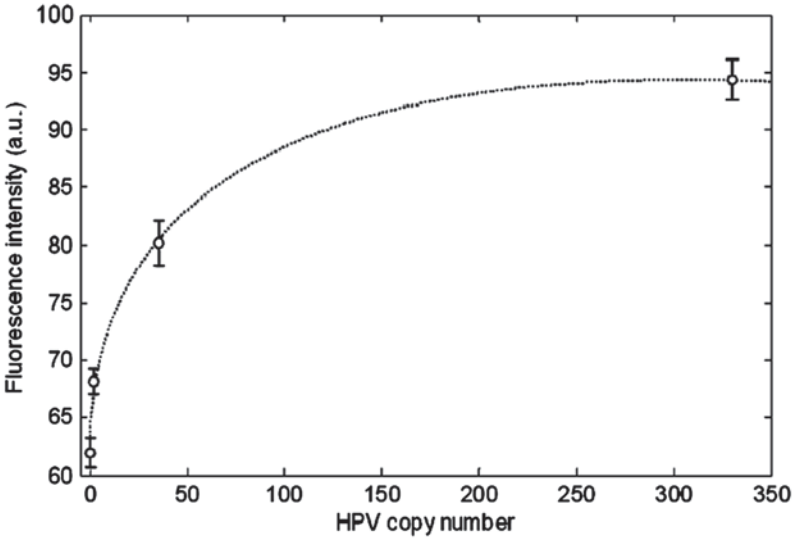


Fig. 13.7 Fluorescence intensity related to the p16INK4A expression level in cervical cancer cell lines plotted against the average HPV copy number present in a cell with the fitted function. The *dotted line* is a fit of the data with the equation $y = a + b\sqrt{x} + cx$, where a = 62.4, b = 3.5 and c = −0.1. (Reproduced from [64])

lation between increasing grade of cervical lesion and staining intensity of $p16^{INK4A}$ was observed [81]. Similarly, Agoff et al. showed that $p16^{INK4A}$ expression correlates with an increasing severity of cervical disease [82]. Murphy et al. showed a strong correlation between HR-HPV positivity and $p16^{INK4A}$ staining pattern [78, 83, 84]. In the study conducted by Wang et al., a correlation between $p16^{INK4A}$ immunostaining and cervical disease severity stratified by HPV status was observed [85].

The results demonstrate a correlation between $p16^{INK4A}$ immunostaining and the presence of HPV in cervical cancer cell lines, supporting the previous studies and implies that $p16^{INK4A}$ is a highly sensitive marker of HPV in cervical cancer cell lines. The hyperbolic-like relationship between the levels of $p16^{INK4A}$ and HPV copy number is typical of the response of a cell to the action of an agonist [86] and in Fig. 13.7, is approximated by a fit with the following function $y = a + b\sqrt{x} + cx$, where a = 62.4, b = 3.5 and c = −0.1. It should be noted that, in Fig. 13.7, the HPV copy number is represented by the average of the range quoted in literature [87, 88] and so error margins in the horizontal axis are potentially very large. However, the sublinear nature of the plot indicates that $p16^{INK4A}$ expression levels are particularly sensitive for low HPV copy number.

FTIR maps were recorded for C33A, SiHa, HeLa and CaSki cells (2–4 maps for each cell line) using the Perkin Elmer Spotlight 400N system. As in Raman spectroscopy, three main biochemical cellular components feature strongly: proteins, lipids and nucleic acids. A detailed examination of the relative intensities of peaks related to vibrations of the cellular components (nucleus and cytoplasm) of the C33A, SiHa, HeLa and CaSki infrared spectra revealed the following tendencies:

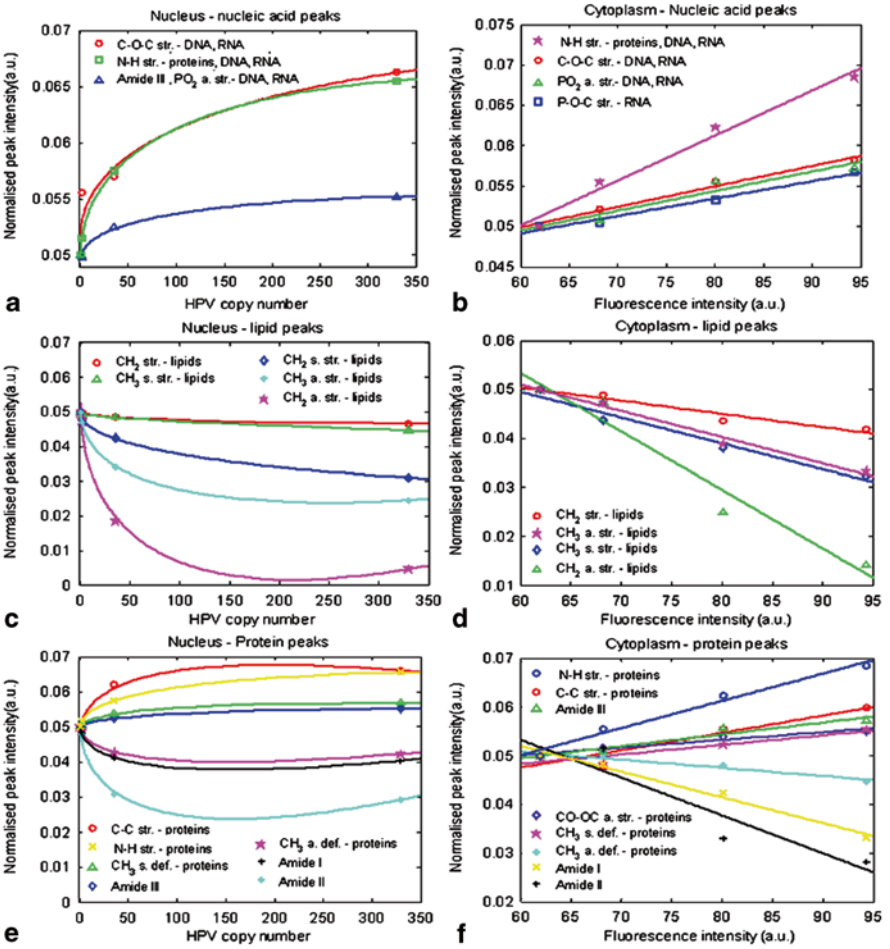


Fig. 13.8 Peak intensity analysis for FTIR spectra of nuclear and cytoplasmic regions of cervical cancer cells. Dependence of peak intensities vs. HPV copy number was fit with $y=a+bIx+cx$ function, while peak intensities vs. fluorescence intensities (p16INK4A expression level) was fit with a linear function, $y=a+bx$. (Reproduced from [64])

- Increase in nucleic acid levels with increasing number of HR-HPV copies in the cell or increasing p16INK4A expression (Fig. 13.8a, d),
- Decrease in lipid levels with increasing number of HR-HPV copies in the cell or increasing p16INK4A expression (Fig. 13.8b, e),
- No tendency for changes in protein levels related to HR-HPV copy number or increasing p16INK4A expression (Fig. 13.8c, f).

The observed increase in nucleic acid levels may be related to an increased number of chromosomes present in HPV infected nuclei. It was reported by Mehés et al. that HPV presence facilitates polyploidisation (increase in chromosome number in a cell

nucleus) in cervical squamous cells [89]. Additionally, it is known that binding of HPV DNA to host DNA disrupts the normal function of the cellular proteins and as a consequence, the host cell accumulates more and more damaged DNA that cannot be repaired [90].

CaSki cells are known to be the most malignant and in the spectra of the cytoplasm and nucleus representing this cell line the highest level of nucleic acids was observed. A similar increase of the nucleic acid related peaks in cancerous cervical cells was noticed in previous studies [91, 92] and is confirmed by these observations. Changes in lipid levels are possibly associated with the disruption of the membrane functionality caused by the virus which influences lipid rafts [93]. A similar behaviour was noted and reported in the previous study on cervical cancer cell lines, wherein both Raman and FTIR (Fig. 13.4) techniques were utilised [63]. Moreover, it was reported previously by Wong et al. that the degree of disorder of methylene chains in membrane lipids increases in cervical neoplastic cells [94]. The absence of any clear relationship between protein levels and HPV copy number is expected, as the presence of HPV in the cell results in functional over- and under-expression of only select biomarker proteins [73], while the level of the other proteins would not be affected. The variation of the spectroscopic features associated with the nuclear lipids, proteins and nucleic acids as a function of the HPV copy number was found to be sublinear and to be well fitted with a similar function to that employed to describe the variation of the $p16^{INK4A}$ levels with HPV copy number in Fig. 13.7. Notably, the dependences of the cytoplasmic lipids, proteins and nucleic acids on the $p16^{INK4A}$ levels, as identified by fluorescence were found to be linear.

In order to further elucidate the multivariate spectral signatures which are specifically related to HPV infection, multivariate regression models were constructed using Partial Least Squares Regression (PLSR) and PLS Jack-knifing was employed as a multivariate feature selection technique [94–96]. The PLSR algorithm seeks to develop a model that relates the spectral data (*X-matrix*) to a series of targets (*Y-matrix*, e.g. concentration of reaction product or analyte) according to the equation $Y = XB + E$, where B is a matrix of regression coefficients and E is the regression residual. The PLS *Jack-knifing* method developed by Westad and Martens [96] was then used to determine the spectral features that were statistically significant at a particular level of confidence using t-testing of the regression coefficients, B.

Figure 13.8 demonstrates that $p16^{INK4A}$ levels as measured by fluorescence intensity are approximately linearly correlated with the univariate spectral features and thus, as PLS is a linear model, these values are employed as targets for the PLS analysis. The model can therefore be applied to the FTIR data to elucidate multivariate signatures which are correlated to $p16^{INK4A}$ expression level, and therefore, in accordance with Fig. 13.7, to HPV infection levels. Once established, these variation patterns can then be applied to unknown samples to screen for the biomarker levels. As $p16^{INK4A}$ was found to be predominantly expressed in cytoplasmic regions of the cells, signals recorded from the cytoplasm were utilised in the analysis. Calibration (based on 40% of the data) and test set performance (based on 60% of the data) are presented in Fig. 13.9a and demonstrate a very good fit to the model. As shown in Fig. 13.9b, the PLS loading exhibits variations originating

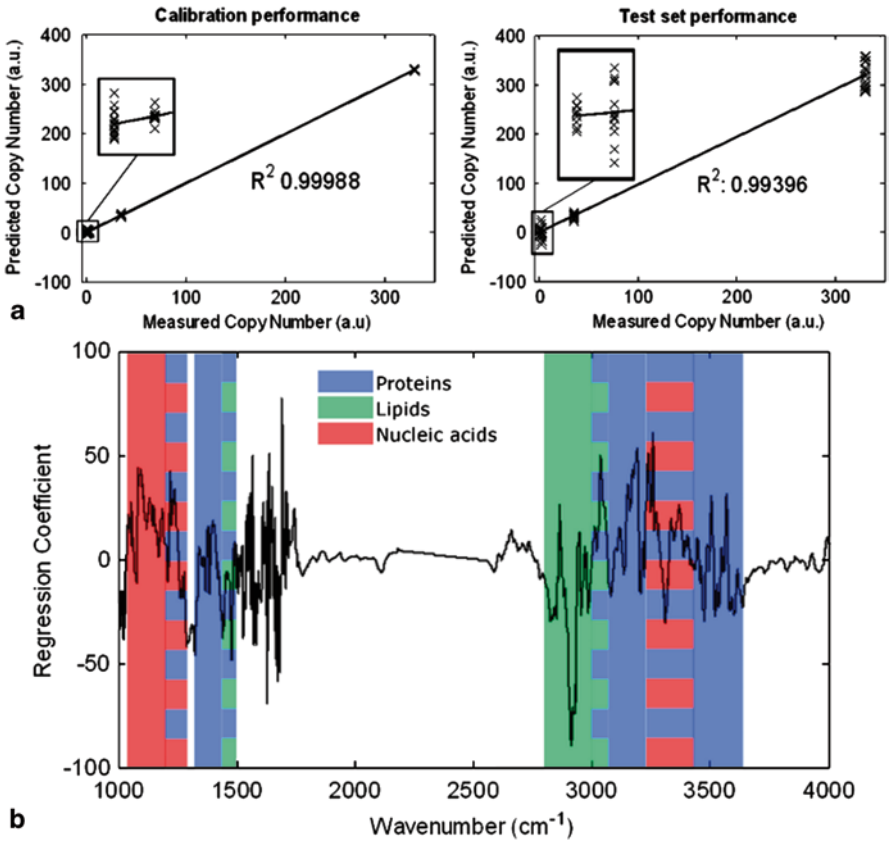


Fig. 13.9 Partial Least Squares analysis results targeting the p16INK4A expression level prediction based on the FTIR spectral features. (Reproduced from [64])

from proteins, primarily as sub-bands of Amide A (3,000–3,700 cm^{-1}) and Amide III (1,230–1,250 cm^{-1}), and are associated with the chain of biochemical disruption in protein regulation caused by HPV presence [97]. Another prominent spectral feature differentiating the cell lines, and present in PC loadings and the PLS loading, is the lipid contribution. Again, the presence of HPV is seen to significantly influence the lipid balance within the cell [93], and thus it can be expected that the vibrations of lipid component groups differentiate cells with various HPV copy number. This differentiation was distinctly manifest in the mean spectra of the cytoplasm and nucleus (Fig. 13.4—enlarged lipid region of 2,800–3,050 cm^{-1}). Furthermore, the DNA and RNA related peaks are exhibited in the range of 1,050–1,250 cm^{-1}, (sym. and asym. PO_2 str. CO str.) and 3,200–3,400 cm^{-1}.

A potential application of vibrational spectroscopy to cervical cancer screening and diagnosis requires a full understanding of the spectral information and its correlation with existing screening and diagnostic methods such as Pap testing or colposcopy. However, the field of the cervical cancer recognition and detection is

still developing with new biomarkers [73] being identified that may be considered as adjuncts to existing cervical cytology and pathology methods. Biomarkers such as $p16^{INK4A}$ have been found to be useful for low-grade lesions [176, 177, 178, 179] making them particularly attractive. Thus, a correlation and explanation of the spectral information with these new biomarkers is a new challenge for biospectroscopists working on future cancer diagnostic systems. The study outlined above demonstrates how multivariate regression techniques can help to associate specific spectral signatures with specific biomarkers, potentially paving the way for "spectro-markers" of disease.

13.4 Analysis of the Effects of Anticancer Agents [98, 99]

The previous sections demonstrate how vibrational spectroscopy, coupled with multivariate statistical techniques can be utilized for high sensitivity and specificity disease screening, but also to analyse the underlying biochemical changes at a cellular level. This analytical capability opens up the possibility of a range of *in vitro* applications.

For the design of new anticancer drugs, an in-depth understanding of the mechanisms underlying their biological effects is required. The *in vitro* study of the interaction of anticancer drugs with mammalian cells is therefore important to elucidate the mechanisms of action of the drug on its biological targets and thereby maximise efficacy and guide development strategies. The prevalence of such studies is likely to increase following the EU Directive (2010/63/EU) on the replacement, reduction and refinement (RRR) of animal experimentation in scientific studies. Vibrational spectroscopy has great potential to explore sub-cellular biochemical structure as exemplified by its use in studies investigating the action of various agents on biological macromolecules as well as their interaction with cancer cells [100–106]. It can provide high content information as it can examine spectral changes in the cell membrane, cytoplasm and nucleus of the target cells simultaneously and highlight the biochemical interaction of external agents with the cell or its physiological response [98, 99]. Validation of the technique of Raman spectroscopy in the quantitative measurement of the biochemical and physiological effects of known chemotherapeutic agents is required, however, before it may be used for studies of novel chemical treatments. The study of the spectroscopic signatures of the biochemical effects of commercially available anticancer drugs may be helpful before application for the evaluation of the novel anticancer agents.

Cisplatin (cis-Diamminedichloroplatinum (II)) is a well established chemotherapeutic agent with a known mode of action. In the nucleus, it binds with DNA forming inter-strand and intra-strand crosslinks which lead to cell cycle arrest and apoptosis [107]. The formation of inter-strand and intra-strand crosslinks between cisplatin and DNA leads to conformational changes of the DNA. To understand the changes caused in DNA conformation due to the adducts formed by cisplatin, the formation of 1,2-GG or 1,3-GTG intra-strand cross-links and their characterization

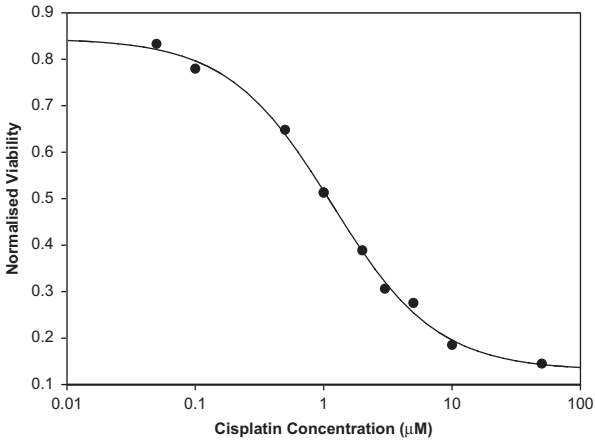


Fig. 13.10 A549 cell viability (measured by MTT absorption) at 96 h after exposure to Cisplatin. Error bars denote the standard error on the mean at each concentration. (Reproduced from [98])

by Raman spectroscopy has been carried out [108] and it was found that 1,3 GTG cross linkages induce more distortion in the B-form DNA as compared to the 1,2-GG complex.

Cancer cell lines, such as the A549 human lung adenocarcinoma cell line provide a good *in vitro* model for lung cancer. The high spatial resolution of Raman microscopy can be employed to identify localized spectroscopic changes in the nucleus, cytoplasm and cell membrane and the responses can be correlated with drug dose, as well as physiological response, as determined by the cytotoxicological assay MTT. Feature selection techniques can then be used to identify regions of the spectrum that are associated with the biochemical effect of exposure to the agent, and with the subsequent cytotoxicological response of the cells.

The MTT assay is a measure of the mitochondrial metabolic activity and can be used initially as a standard against which to map the spectroscopic response of the cells. Maximum cytotoxicity was observed after 96 h exposure and thus this time period was employed for all measurements. The cytotoxicity induced in A549 cells due to the exposure to cisplatin for 96 h, as determined by the MTT assay, is shown in Fig. 13.10, where the level of viability in each sample was normalised to that in the control sample. Due to the action of the drug, the mitochondrial activity decreases monotonically which in turn leads to a decrease in cell viability. The Inhibitory Concentration (IC_{50}) value was derived from the data by a fit of $f(x) = \min + (\max - \min)/(1 + (x/IC_{50})\text{^}n)$ and found to be 1.2 ± 0.2 μM, which is consistent with the literature. For A549 cells exposed to test drug concentration (TDC) for 72 h, IC_{50} values of 3.59 and 2.2 μM are reported elsewhere with MTT and ATP assays respectively [109]. Also, Cordes et al. have determined an IC_{50} value of 2.0 μM for cisplatin (0.1–50 μM) exposed A549 cells using the colony formation assay [110].

While some indication of the localised biochemical impact of the external agent can be gleaned from the mean spectra of the cellular regions [98], the spectral changes are multivariate and may have complex exposure dependence. They

include the fingerprint of the interaction of the cisplatin with the nuclear DNA, but also evident are a multitude of other changes associated with exposure and changes in cellular physiology. In an effort to identify the most important spectral features, feature selection was performed using the PLS—Jack knifing analysis described earlier. The regression co-efficients were obtained forseparate regressions against cisplatin concentrations and against cell viability with the nuclear spectral data. To elucidate spectral changes that are most statistically significant with respect to regression against either endpoint, a t-test ($p < 0.001$) was applied on the PLS-Jack knifing results. The spectral features identified by this process are highlighted by vertical bands in Fig. 13.11.

The PLS regression of the nuclear data against cisplatin concentration (Fig. 13.11a), identifies many features but, for the purpose of this discussion, only the changes which are related to the known action of cisplatin are highlighted. Features identified include the thymine (737–742 cm^{-1}, 1,589–1,595 cm^{-1}), guanine (1,548–1,553 cm^{-1}) and adenine (1,578 cm^{-1}, 1,604–1,611 cm^{-1}) bands. The regression co-efficients of guanine and thymine are stronger than that of adenine, indicating that cisplatin binds preferably with guanine as compared to adenine [111]. Another interesting observation is that the bands at (877–883 cm^{-1}) and at (899–900 cm^{-1}), both assigned to deoxyribose ring breathing, are significantly changed, indicating a change in the conformation of B-form DNA [112–114]. Changes in the Raman bands of the proteins are also highlighted, as indicated by the regression co-efficients which can be assigned to tryptophan ring breathing (761–765 cm^{-1}), (889–891 cm^{-1}), C-N stretching (1,128–1,132 cm^{-1}), C-H deformation (1,315–1,316 cm^{-1}, 1,351–1,354 cm^{-1}), amide-I (1,668–1,679 cm^{-1}). These changes may indicate some structural changes in the nuclear proteins. The spectroscopic features selected by the PLS regression against drug concentration are consistent with the known direct interaction of the drug with the DNA and thus are identified as spectroscopic markers of the direct chemical interaction of the drug.

The regression against cell viability (Fig. 13.11b), shows the features which are correlated with the change in cell physiology. These include the conformational changes of the DNA B-form (832–833 cm^{-1}) and DNA A-form (798–813 cm^{-1}). These changes in the Raman bands of the DNA can be taken as markers for the conformational changes in the DNA caused by cisplatin binding to the guanine and thymine DNA bases (Vrana et al. 2007) and also confirm the observations from the mean difference spectra. Changes related to protein secondary and tertiary structures are indicated by the selection of amide-III random coil (1,225–1,228 cm^{-1}), amide-III β-sheet (1250–1259), amide-III α helix, tryptophan ring breathing (711–721 cm^{-1}, 758–763 cm^{-1}, 767–772 cm^{-1}) and phenyl ring stretching modes (1,202–1,210 cm^{-1}) together with the C-C skeletal vibrations (935–937 cm^{-1}, 1,159–1,163 cm^{-1}), C-N stretching vibrations (1,065–1,068, 1,073–1,075) and C-H bending (1,178–1,182 cm^{-1}) (1,185–1,186 cm^{-1}). These features suggest changes in the Raman bands of proteins and are indicative of structural alterations associated with protein activation and denaturation during cellular apoptosis and nucleic acid repair responses.

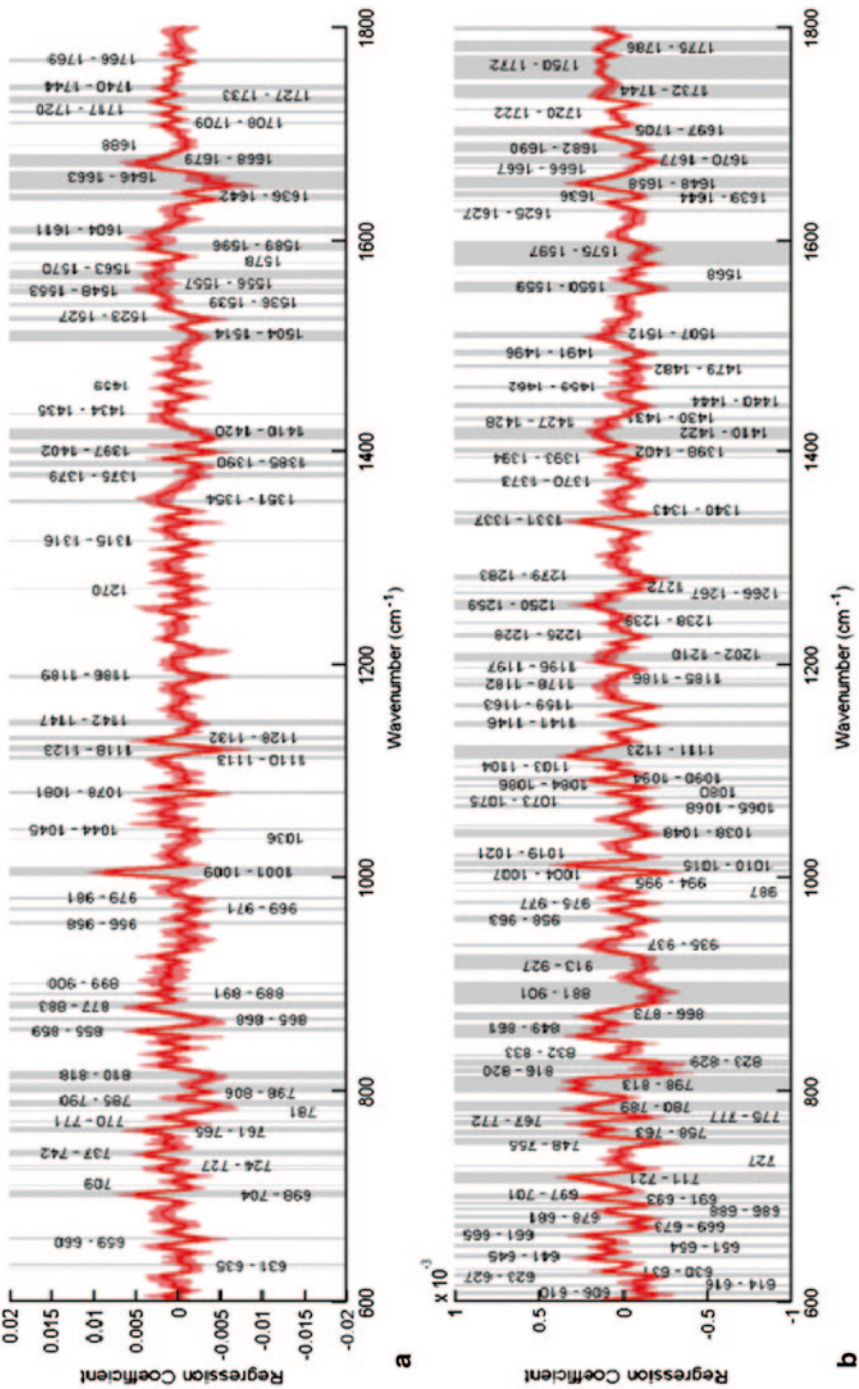


Fig. 13.11 Feature selection in PLSR of Raman spectra of the A549 nucleus against **a** cisplatin concentrations, and **b** against cellular viability. The *shaded areas* denote statistically significant features selected using PLS Jack-knifing with a t-test ($p < 0.001$) of regression coefficients (shown in *red* with associated standard error). (Reproduced from [98])

Comparing the regression co-efficients obtained via spectral regression against cisplatin concentration (which highlight the primary biochemical effect of the action of the agent) with those obtained via regression against cell viability (which highlight biochemical markers of the change in the viability of the cell), very different spectral features are selected as being statistically significant and thus allow the differentiation of the chemotherapeutic response. The spectral features of thymine (737–742 cm^{-1}, 1,589–1,595 cm^{-1}), guanine (1,548–1,553 cm^{-1}) and adenine (1,578 cm^{-1}, 1,604–1,611 cm^{-1}) emerge as prominent as a result of regression against the concentration of the chemotherapeutic agent and are attributable to direct binding of the cisplatin molecule with these DNA bases. In addition, changes related to proteins feature in the regression against the concentration of the chemotherapeutic agent are significantly different than those changes of protein associated bands which appear as a result of regression against the cellular viability. The changes in the features of the proteins due to the chemical effect are attributable to structural alterations in nuclear histone proteins due to cisplatin interactions with the cellular DNA. These features are not selected as being statistically significant by the analysis for regression against the viability endpoint, suggesting that they are markers of the biochemical interaction between cisplatin and DNA. Contrastingly, the conformational changes to the DNA A-form (798–813 cm^{-1}) and DNA B-form (832–833 cm^{-1}), among the changes associated with the protein features, especially the amide-III random coil (1,225–1,228 cm^{-1}), amide-III β-sheet (1,250–1,259), amide-III α-helix, among others, are selected as being statistically significant in the regression against cellular viability, suggesting that they are markers of this physiological endpoint in the chemotherapeutic response. Overall, Fig. 13.11 demonstrates that the Raman spectrum of the cell may be resolved into distinct groups of statistically significant spectral markers for distinct types of effect. A similar analysis of the spectral changes as a result of the action of the chemotherapeutic agent in the cytoplasm and cell membrane can be performed [99]. Raman microspectroscopy can therefore clearly fingerprint the effects of the interaction of the chemotherapeutic agent with DNA, and identify that fingerprint within the sub cellular spectra of exposed cells.

PLSR models of A549 nuclear spectra versus concentration and reference viability measurements from the MTT assay can then be constructed to determine the ability of the spectral data to predict the level of exposure to cisplatin and the associated level of physiological effect from Raman spectra. All spectra were compiled into a matrix, and were randomly sorted. A total of 60 % of the spectra were used to train the PLSR model and 40 % of the total was retained as an independent test set to assess the performance of the model in predicting the level of exposure, and the cellular viability, with unseen data. Leave-one out cross validation with the calibration set was used to determine the optimal model complexity for use in testing [115]. This process was performed on fifty separate occasions, with randomization of the data matrix and splitting of the data on each occasion to prevent data bias [116]. The results of prediction of levels of cellular viability is shown in Fig. 13.12, for the example of the nuclear region, where the values of the root mean square error of calibration (RMSEC) and the root mean square error of prediction

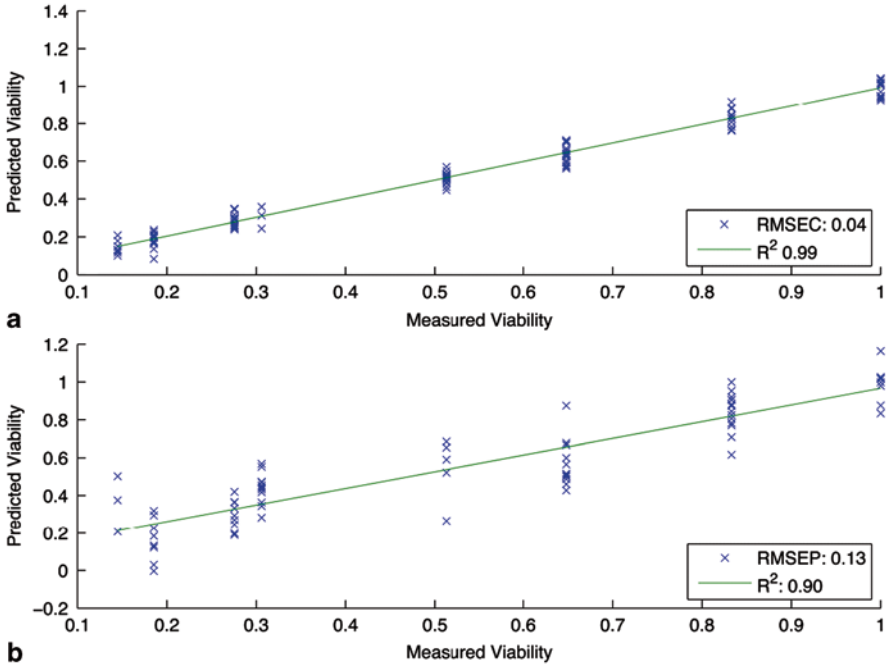


Fig. 13.12 Calibration (**a**), and test set performance (**b**), in prediction with PLSR of the cell viability using Raman spectra of the nucleus of A549 cells. Root mean squared errors of calibration (RMSEC) and prediction (RMSEP) are in units of cell viability, which themselves are normalized to the level in the control. (Reproduced from [98])

(RMSEP) denote the prediction uncertainty. The mean values of the RMSEC and RMSEP for the PLSR for prediction of the level of exposure to cisplatin were found to be 1.67 μM (0.45) and 3.41 μM (0.41) respectively, where the figures in brackets denote the standard deviations of these estimates. The associated values of RMSEC and RMSEP for PLSR against normalized cellular viability were 0.05 (0.008) and 0.11 (0.011) respectively. The RMSEP for prediction of cisplatin concentration is therefore 6.8% over the full scale range (0–50 μM), and 11% over the full scale range of viability (from 0 to 1). The proportionally higher observed errors of prediction of cell viability may be due to the development of cellular resistance to cisplatin over the exposure period in the low dose groups [117, 118], which contributes to a non-linearity in the variation in viability with concentration that is difficult for the PLSR algorithm to model.

The study demonstrates the potential of commercially available benchtop Raman microscopy for the examination and prediction of chemotherapeutic responses. In general, for subcellular resolution, Raman spectroscopy is chosen over infrared spectroscopy due to the intrinsically higher spatial resolution. Furthermore, the relative insensitivity to water renders live cell imaging of real time exposures a possibility [119]. Nonlinear optical imaging techniques [120] promise even higher

spatial resolution in the future, but at present such short pulse systems cannot be considered benchtop techniques.

Cisplatin, a DNA major groove binder [121], was chosen as a model as its mechanism of action is well characterised. To fully validate the technique, the study should be extended to other chemotherapeutic agents of known interaction mechanisms, such as tallimustine, a known DNA minor groove binder, and ethidium bromide, a known DNA intercalator. In this way the spectroscopic signatures of their interaction can be established and subsequently can act as a database against which novel chemotherapeutic agents can be evaluated. Ultimately, the techniques should be validated in tumour tissue.

13.5 Live Cell Imaging [122]

Raman microspectroscopy presents several advantages for the study of live cells, combining molecular analysis with optical imaging. In comparison to infrared spectroscopic analysis, the weak contribution from water offers the possibility to study the cells in an aqueous environment and thus to keep them alive for the duration of the measurement [101]. The specific information contained in the Raman spectra provides a signature of the samples studied, which can be related to molecular content or changes to the physiology as a result of external stimuli [123, 124]. The spatial resolution is of the order of 0.5-2 μm, providing access to the subcellular organisation of the cells at a molecular level [125, 126].

Previous studies have demonstrated the capacity of Raman spectroscopy to map single cells and their subcellular structures. The identification of the sub-cellular structures such as the nucleus, cytoplasm and mitochondria can be achieved [127] and chemical maps can be constructed giving information about the distribution of differentiating molecules such as proteins [128]. However, these have been performed on chemically fixed cells and it has since been demonstrated that chemical fixation induces significant changes to the biochemistry of the cell [129, 130]. Ultimately, if subtle changes in metabolism due to for example chemotherapeutic agents, toxicants, or radiation are to be discerned, live cells should be studied.

Over the last decade, new models such as collagen gels and more complex multicomponent systems such as Matrigel have emerged, providing the cells a matrix which more accurately reproduces the extra cellular matrix (ECM) [131–133]. It has been demonstrated in various studies that the behaviour of cells cultured in such an environment differs considerably compared to those cultured on traditional 2D surfaces [134]. In terms of cellular viability and spectral quality, collagen gels have been recently demonstrated to be a substantially improved substrate for live cells analysis using Raman spectroscopy [122]. The gels provide a 3-D matrix for cell growth, although it should be noted that the cells are cultured on the surface of the matrix. As measured using the Alamar Blue cytotoxicity assay, in comparison to uncoated quartz, a human lung cell line exhibited a 39% increase in viability in collagen gels whereas as dermal cell line seemed to be more affected by the differ-

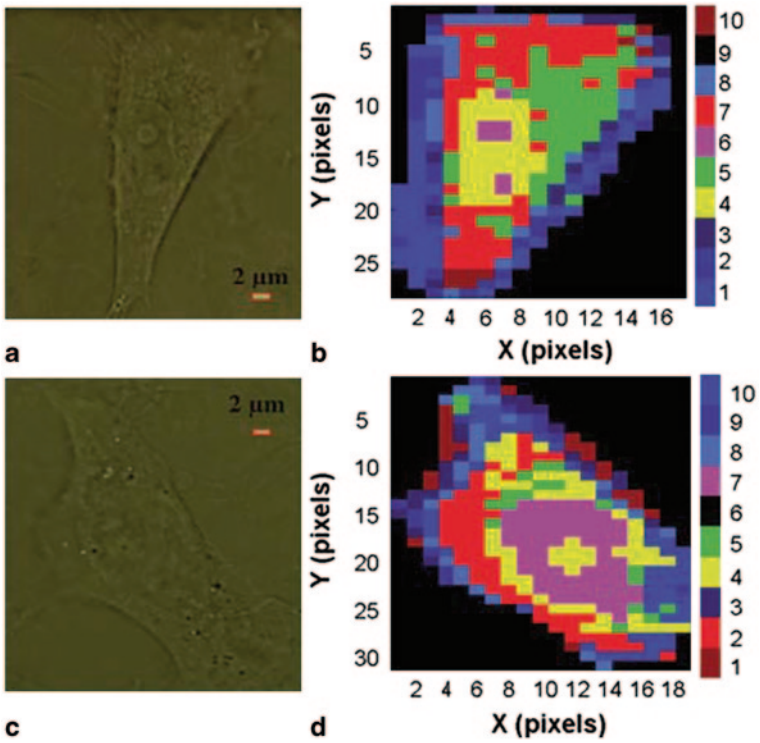


Fig. 13.13 **a** and **c** are examples of images taken from the Labram HR 800 equipped with x100 immersion objective respectively from a A549 cell and a HaCaT cell. **b** and **d** are the corresponding false colors images obtained after K-means cluster analysis of Raman maps. Each pixel of the Raman map has been attributed to a specific cluster. The cluster numbers are indicated in the colour code to the right of the figure. (Reproduced from [122])

ent substrates and a 285% increase in viability was observed. Furthermore, after polymerization, the density of the collagen is too low to give any contribution in the spectra recorded [119]. Also no contribution from the medium used for the cell culture is observable in the spectra recorded which allows the cells to be kept alive for prolonged periods and spectroscopically monitored, effectively in real time.

After growing A549 and HaCaT (immortalised human keratinocyte) cells for 24 h on collagen gels, the medium was substituted by NaCl before mapping, as the phenol red of the cell growth medium limits the visibility of the cells [119]. Examples of photographs obtained using a Horiba Jobin—Yvon Labram HR800 equipped with a x100 immersion objective are presented in Fig. 13.13a, c. The different structures of the cells can be identified; the cell membrane is clearly defined for both cell lines and the nucleus is located in the centre part of the cells. The cytoplasm appears smaller in the case of the HaCaT cell. The main difference between the two cell lines is the organisation in the nucleus. When observing the A549 cells, the nucleoli are clearly visible inside the nucleus but in the case of the HaCaT cell,

no well defined shapes can be observed under these conditions. Another observation is the presence of 2 different organelles in the cytoplasm of the A549 cell. This observation is not surprising as the cytoplasm of the cells has a complex organisation and different structures can be found such as the actin or microtubules [135, 136], endoplasmic reticulum [136, 137], the mitochondria [138] or the Golgi apparatus [139, 140].

Raman spectral maps were performed in immersion using a Horiba Jobin Yvon LabRam HR800 using a 785 nm laser as source. The step between two successive measurements was set to 1.5 μm and the backscattered Raman signal was integrated for 10 s over the spectral ranges from 400 to 1,800 cm^{-1} and accumulated twice to improve the signal to noise ratio. In order to visualise the subcellular structures in the spectral maps recorded, K-means clustering analysis was employed.

K-means clustering analysis is one of the simplest unsupervised learning algorithms used for spectral image analysis. It groups the spectra according to their similarity, forming clusters, each one representing regions of the image with identical molecular properties [141]. The distribution of chemical similarity can then be visualised across the sample image. The number of clusters (k) has to be determined a priori by the operator before initiation of the classification of the data set. K centroids are defined, ideally as far as possible from each other, and then each point belonging to a data set is associated to the nearest centroid. When all the points have been associated with a centroid, the initial grouping is done. The second step consists of the calculation of new centroids as barycentres of the clusters resulting from the previous step. A new grouping is implemented between the same data points and the new centroids. These operations are repeated until convergence is reached and there is no further movement of the centroids. Finally k clusters are determined, each containing the most similar spectra from the image. From here, colours can be attributed to each cluster and false colour maps can be constructed to visualise the organisation of the clusters in the original image.

K-means clustering analysis is currently commonly used for the analysis of tissue sections by infrared and Raman spectroscopy, especially for the detection of the different structures present in the samples studied and the identification of tumoral regions [142–144]. K-means clustering can also be employed for the study of individual cells by Raman spectroscopy but only a few studies have elucidated the subcellular organisation and they are usually based on fixed cells grown on a non optimal substrate such a quartz window or equivalent [127, 145]. The effects of cell fixation on the cellular composition has been already investigated [129, 130] and thus it is desirable to be able to obtain the Raman signature from live cells, avoiding any structural and molecular modification which may occur as a result of the chemical processing. Moreover, the use of collagen gels enables extended measurement periods and thus more detailed mapping of live cells. In contrast to quartz substrates which have a toxic effect necessitating fixation of the cells for prolonged measurement times, collagen gels offer a better adhesion of the cells to the substrate, enhancing the cell viability.

Within the cells, it may be predicted that distinct biochemical regions can be identified, corresponding to, for example, nucleus, nucleoli, cytoplasm and cell

membrane. However, in analysing the spectral response of the cells, it is observed that the spectra recorded from the edge of the cells contain a high variability due to the transition from the membrane to the substrate, resulting in quite noisy spectra. When the K-means analysis is performed, the variability between these spectra can interfere with the analysis, resulting in the creation of new clusters. Thus, in order to best visualise the different structures existing inside the cells the numbers, the number of clusters has to be increased to 10 groups, which produced the best reproducibility. In this way, the variability in the spectra obtained at the edge of the cells are contained in different clusters and do not interfere with the identifications of the different structures present within the cells. The initiation step was selected randomly for each iteration and it was seen to have no effect on the final results of analysis. After analysis, the spectra were gathered in the 10 different groups according to their similarities and a colour (arbitrarily) attributed to each cluster. False colour maps were then constructed representing the partition of the different clusters in the cells. The resulting images are presented in Fig. 13.13b, d. A correspondence between the organisation of the clusters and the structure present in the cells can be seen. In Fig. 13.13b, specific clusters relating to the cytoplasm, nucleus and nucleoli appear. The specificity of the information contained in the spectra recorded allows discrimination of the different cellular compartments using K-means analysis. The sensitivity of the analysis furthermore allows visualization of the two different compartments present in the cytoplasm of the A549 cell and identifies a specific cluster related to the nucleoli in the HaCaT cell, although they are not clearly seen in the visible image.

K-means clustering analysis appears to be a powerful tool to identify the different structures present in the cells. The false colour maps give a representation of the partition of the different clusters but do not provide any information about the chemical composition of the samples. This biochemical information and the variations between the different clusters can be visualised through the mean spectra of the clusters. Figure 13.14 presents the average spectra obtained from selected clusters of the K-means analysis performed on an A549 cell (Fig. 13.13b). (A: nucleus from the cluster number 4, B: nucleolus from the cluster number 6, C: cytoplasm from the cluster number 5 and D: cytoplasm from the cluster number 7). The other clusters correspond to the spectra recorded from the collagen gel and those located at the edge between the cell and the substrate (data not shown). Variations can be observed in different spectral windows according to the different clusters (highlighted in grey in Fig. 13.14).

The characterisation of live cells on 2D substrates is well documented [123, 124, 147–149] and, based on literature; assignments of the different peaks can be performed. For example, the variations in the spectral range 718–725 cm^{-1} are related to the lipids and nucleic acids, between 815 and 830 cm^{-1} to RNA and proteins, from 1,088 to 1,102 cm^{-1} to the nucleic acids, the lipids and carbohydrates, from 1,243–1,266 cm^{-1} to the nucleic acids, the proteins and lipids, from 1,296–1,322 cm^{-1} to the lipids and from 1,315–1,348 cm^{-1} to the nucleic acids and proteins. In analysing the cluster average spectra, it is important to note that, with a ~1.5 μm spot size, while the nucleus can be specifically targeted, the overlying cellular membrane and

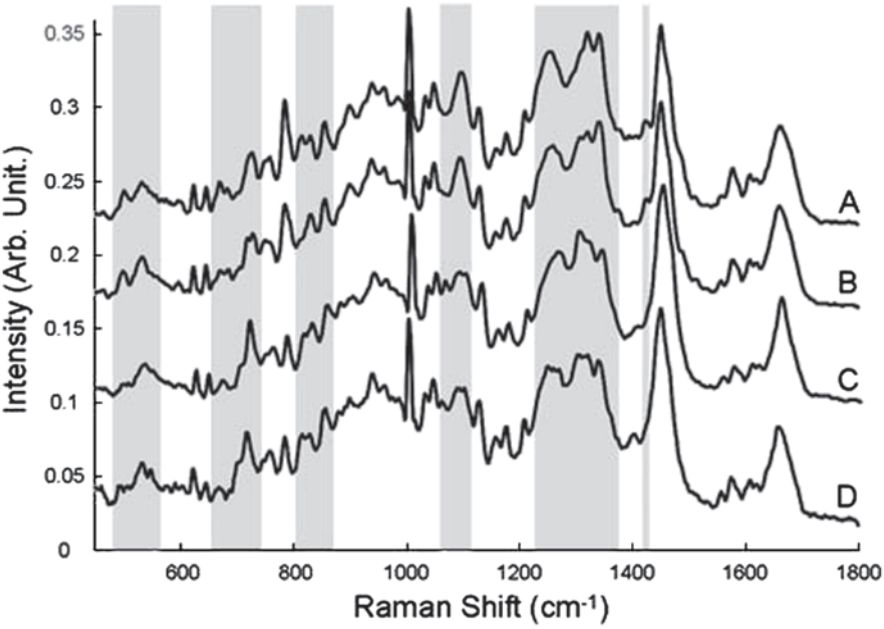


Fig. 13.14 Average spectra obtained from different cluster after K- means clustering analysis. *A* nucleus, *B* nucleolus, *C* cytoplasm cluster 1 and *D* cytoplasm cluster 2. (Reproduced from [122])

cytoplasm will also contribute to the spectrum acquired. Similarly, the spectrum of the cytoplasm will contain contributions from the overlying cellular membrane and potentially also from subcellular organelles such as mitochondria, lysosomes, etc. Across, the spectral range, the variations between the different regions are thus subtle. However, the regions do clearly cluster by K-means analysis, and for example the nucleus and nucleoli are clearly differentiated from the other spectra by for example the peak at 1,423 cm^{-1}, which is completely absent in the cytoplasmic spectra.

The mean spectra obtained from each cluster provide information about the spectral ranges which differentiate the sub-structures of the cells. However, it is difficult to determine precisely the differentiating features and to profile the biochemical structure of the cellular compartment analysed. Using K-means clustering analysis, the different structures of the cells have been identified and spectra associated with the nucleus, nucleoli and cytoplasm can now be extracted from the map and analysed independently to highlight the modification in the spectral signatures. For this purpose, principal component analysis (PCA) was employed, for the case of the A549 cells for illustration purposes [150]. Similar results are obtainable for the HaCaT cell line. Figure 13.15 is a representation of the plots obtained from the spectra extracted from the A549 cell map. Each spectrum has been considered as a single unit for the analysis and the difference compared to K-means analysis is that PCA is a totally unsupervised method. The result is a representation of the

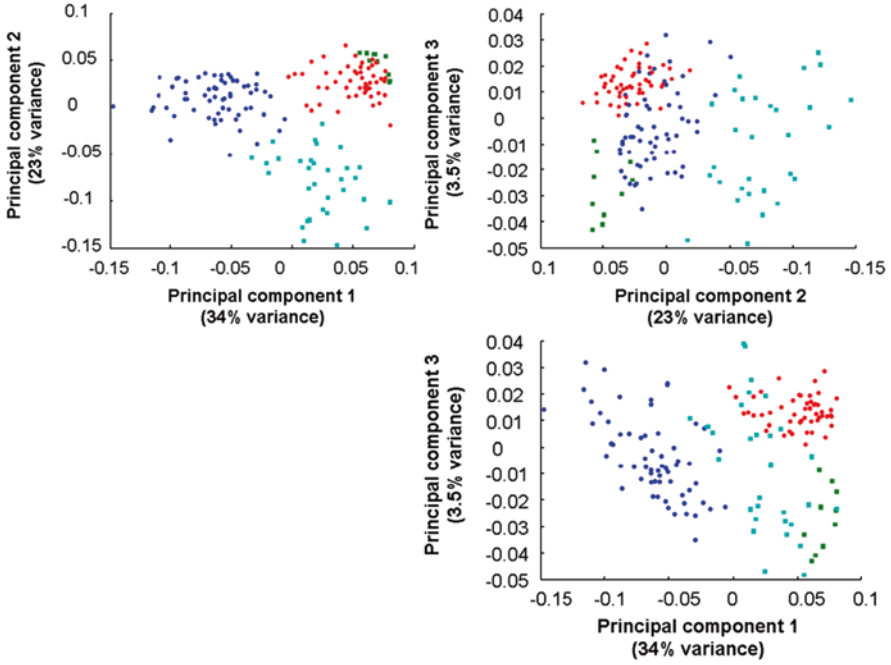


Fig. 13.15 Plots of the first 3 principal components after PCA performed on Raman spectra recorded from A549 cells grown on collagen gels. The spectra have been selected for the cluster corresponding to the nucleus (*red*), nucleolus (*green*) and cytoplasm (cluster 1: *dark blue*; cluster 2: *light blue*). (Reproduced from [122])

discrimination between the different spectra. Figure 13.15a is a plot of PC 1 versus PC 2 for each spectrum. The spectra from the nucleus and nucleoli form one group and are not discriminated by these principal components. Nevertheless, PC 1, which represents 34% of the explained variance, gives a good discrimination between the spectra from compartment 1 of the cytoplasm from the other spectra (this can also be seen in Fig. 13.15c) whereas PC 2, which represents 23% of the variance, allows the discrimination between compartment 2 of the cytoplasm and the other spectra. The second plot (Fig. 13.15b) represents the loadings of PC 2 versus PC 3. PC 2 again gives a clear discrimination between the spectra from compartment 2 of the cytoplasm and the other spectra, as seen in Fig. 13.15a. PC 3, which represents 3.5% of the explained variance, discriminates between the spectra from the nucleus and those from the nucleoli. The same observation can be made in Fig. 13.15c. PCA thus provides a clear discrimination between the different groups of spectra analysed. The results are not surprising, as the spectra have been extracted from different clusters obtained after the K-means analysis and no additional information about the chemical composition of the samples is given in the plots. Nevertheless, the observation of the plots is essential to identify the principal component involved in the discrimination between the different groups, as for each principal component a corresponding loadings is given after analysis.

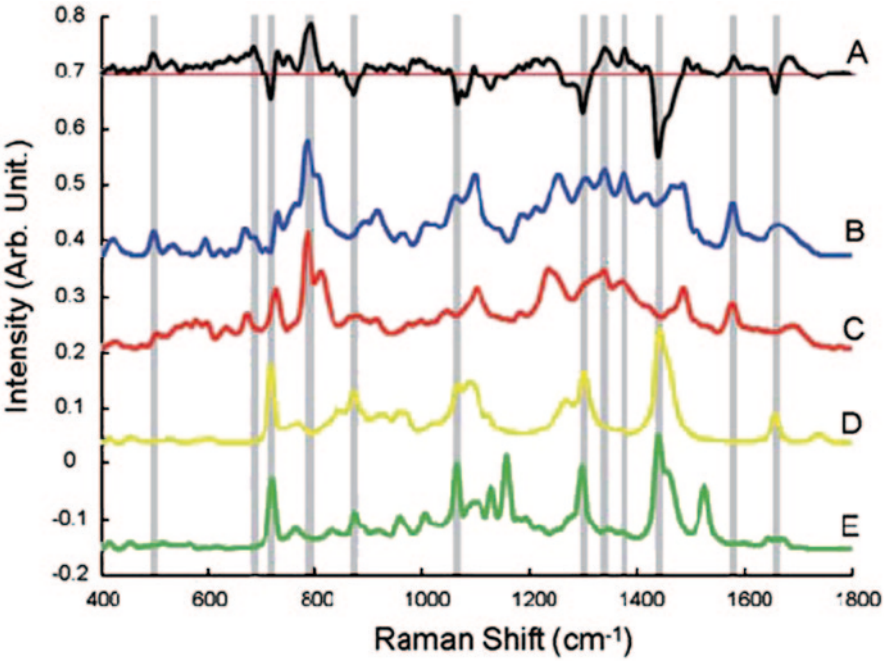


Fig. 13.16 Plot of the loadings number 1 (*A*) corresponding to the first principal component resulting of the PCA analysis. This loading has been compared with spectra recorded from different compounds: DNA (*B*) RNA (*C*) L-α-Phosphatidylcholine (*D*) and Sphingomyelin (*E*). (Reproduced from [122])

The loading is a representation of the variability described by a principal component as a function of wavenumber of the spectra. Figure 13.16a shows the loading of PC1 obtained from the PCA analysis based on the A549 cell map which allows visualisation of which part of the spectrum is responsible for the discrimination between compartment 1 of the cytoplasm and the other spectra, observed in the plot of Fig. 13.15. The loading is a combination of different peaks with different intensities, the solid red line representing the zero level. It appears that the loading is composed of different sharp peaks, both positive and negative, and correspond to increased or decreased contributions of specific molecular components to the Raman spectra. The spectra in Fig. 13.16 A, B, C, D, E, offset for clarity, are from DNA, RNA, L-α-Phosphatidylcholine and Sphingomyelin respectively. DNA and RNA are abundantly present, especially in the nucleus, phosphatidylcholine is a major component of biological membranes and sphingomyelin is also found in animal cell membranes. To demonstrate further the origin of the spectral variations, and their clustering using PCA, the 12 peaks with highest variability have been highlighted in Fig. 13.16a. Each of them has a correspondence with one or more of the four spectra from compounds tested. For example, the peak at 791 cm^{-1} is prominent in the spectra of DNA and RNA. In Fig. 13.14, it can be seen that while all average spectra exhibit a peak in this region, it is significantly stronger in the

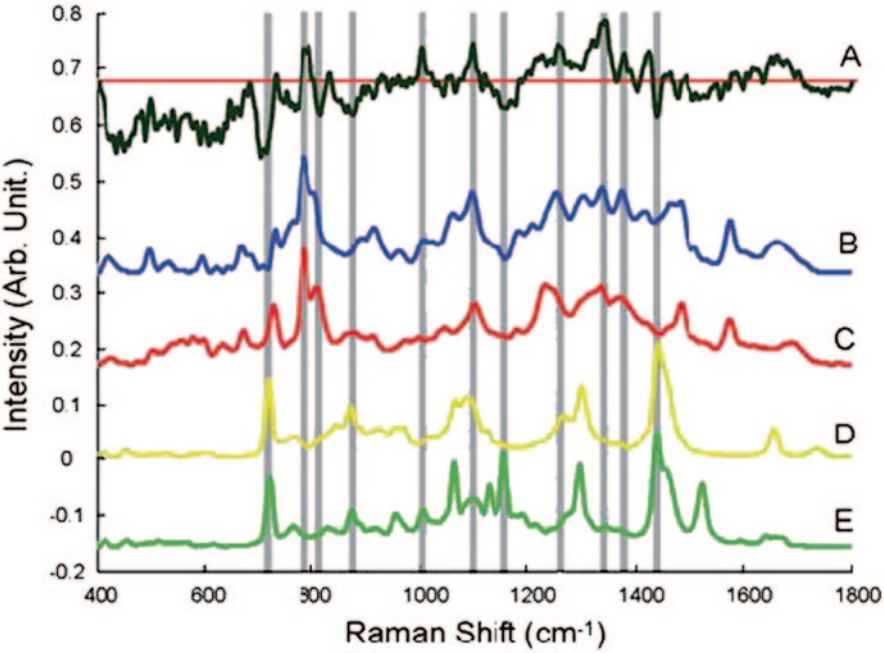


Fig. 13.17 Plot of the loadings number 2 (*A*) corresponding to the second principal component resulting of the PCA analysis. This loading has been compared with spectra recorded from different compounds: DNA (*B*), RNA (*C*), L-α-Phosphatidylcholine (*D*) and Sphingomyelin (*E*). (Reproduced from [122])

spectra of the nucleus and nucleoli, thus giving rise to a strong positive contribution to PC1, which differentiates between nucleus/nucleoli and cytoplasm 1. The strongest variation, however, is the negative peak at 1,439 cm^{-1}, which is prominent in the spectra of the L-α-Phosphatidylcholine and Sphingomyelin indicating significant variations in lipidic content. The loading of PC2, plotted in Fig. 13.17a is the basis of the discrimination between compartment number 2 of the cytoplasm and the nucleus and nucleoli (Fig. 13.15a). Once again the comparison with the spectra of pure compounds highlights peaks matching with the loading. The highest variability can be seen at 699 and 1,436 cm^{-1}, both attributed to the lipid content, and at 1,342 cm^{-1}, related to the DNA and RNA. The loading of PC3 (Fig. 13.18a) represents the variability between the nucleus and the nucleolus (Fig. 13.15b, c). Many peaks can be identified as related to the DNA and RNA, especially at 651, 813, 1,091, 1,482 and 1,572 cm^{-1}, showing that the main difference between the spectroscopic signatures of the nucleus and nucleolus is the DNA/RNA content.

The loadings of the principal components provide a multivariate profile of the biomolecular variations within the cell. While a detailed analysis is not undertaken here, it is clear that a wealth of information can be derived from the Raman analysis of single cells. In contrast to other studies, the work presented here is conducted on live cells rather than chemically preserved, and moreover the details of the spectra

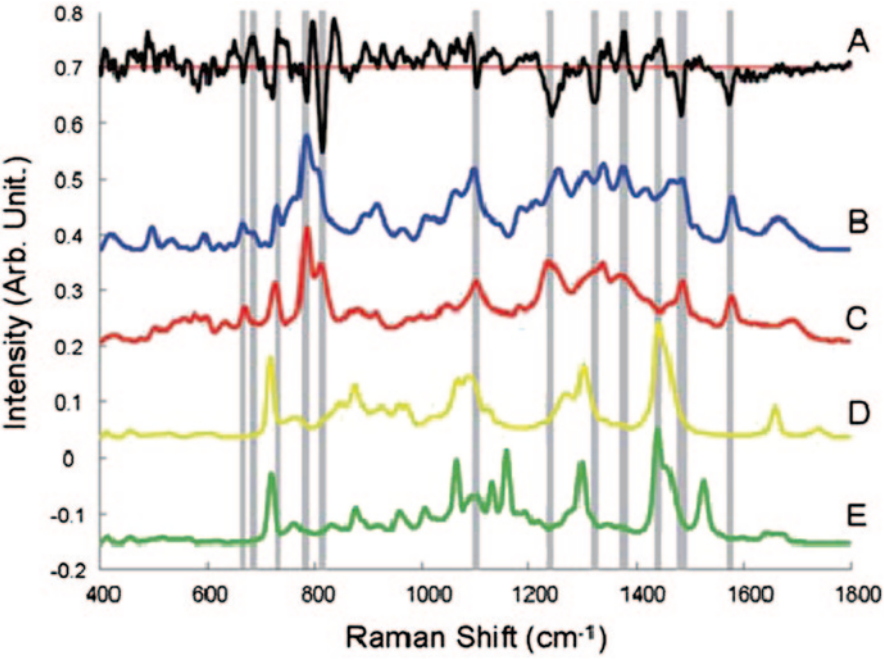


Fig. 13.18 Plot of the loadings number 3 (*A*) corresponding to the third principal component resulting of the PCA analysis. This loading has been compared with spectra recorded from different compounds: DNA (*B*) RNA (*C*) L-α-Phosphatidylcholine (*D*) and Sphingomyelin (*E*). (Reproduced from [122])

are greatly enhanced by the minimal need for data-preprocessing and prolonged measurement periods. These conditions are facilitated by the use of collagen 3-D matrices for cell growth and measurement.

13.6 NanoParticles in Cells [151]

The potential of vibrational, particularly Raman, spectroscopy as a realtime label free subcellular analytical technique opens up a range of applications, as for example in the field of nanotoxicology and nanomedicine. One of the major concerns regarding the possible toxic effects of nanoparticles is the capacity of these materials to penetrate cells and possibly translocate to other cells around the body. Conventional *in vitro* cytotoxicity assays, such as Alamar Blue, Neutral Red, MTT, etc, provide indications of impact on cell proliferation, viability, metabolic activity, lysosomal and mitochondrial activity. However, they fail to provide a verification of nanoparticle internalisation or elucidate intracellular trafficking mechanisms and subcellular distribution within cells. Indeed, false positive results have been demonstrated due to the extracellular interaction of the nanoparticles with the cell culture

medium *in vitro*, and with the cytotoxic assays themselves [152, 152]. The underlying principles of internalisation and distribution of nanoparticles within cells and any associated toxicity to humans are still relatively poorly understood. However, there are numerous studies that have demonstrated the ability of nanoparticles to cross membranes and internalise within many different cell types [153–155]. In order to assess potential risks, but also benefits in terms of nanomedicine, a greater understanding of migration of nanoparticles intra- and inter-cellularly and essentially from one part of the body to another, along with information on the translocation of nanoparticles within organelles, is urgently required.

Among the main challenges faced by nanotoxicologists are the detection and identification of nanoparticles that have crossed the cell membrane and monitoring of their intracellular trafficking. Imaging cells exposed to fluorescently labelled nanoparticles using confocal fluorescence microscopy is one of the most common ways in which to ensure nanoparticles can be tracked and monitored as they enter and localise within cells [156–158]. However, not all nanoparticles can be easily fluorescently labelled. Furthermore, there have been reports that labelled nanoparticles can release the dye into the surrounding biological environment, and so the distribution of fluorescence within the cell does not necessarily represent the presence or subcellular distribution of the nanoparticles [159–161]. Furthermore, it is not clear that the transport mechanisms of smaller nanoparticles, fluorescently labelled with anionic moieties, are the same as their unlabelled counterparts. Transmission Electron Microscopy can also be employed to visualise nanoparticles within cells, but significant sample processing (fixing and ultramicrotoming) is required and only particles with sufficient electronic contrast to the cellular environment can be visualized [162, 163]. Thus, there is a requirement for a sensitive technique to localise and identify nanoparticles internalised in cells, ideally based on their chemical composition, rather than fluorescence labels or electronic contrast. Identification of their local environment (e.g. endosomes, lysosomes, mitochondria) could aid in elucidating their intracellular trafficking and interaction mechanisms, and the resulting changes in cellular metabolism.

Raman spectroscopy potentially offers a label free probe of nanoparticles within cells, which can potentially analyse their local environment, and ultimately changes in the cellular metabolism which can be correlated with cytotoxic responses [164], oxidative stress, or inflammation. Kneipp et al. have previously demonstrated the use of Surface Enhanced Raman Scattering from gold nanoparticles and nanoaggregates to probe the environment of the subcellular compartments through which they are trafficked [165, 166]. They have also demonstrated the use of molecular labelled nanoparticles as more specific probes of the local environment [167–169]. However, the uptake rates and mechanisms as well as the subsequent trafficking may be specific to the nanoparticle type, size and surface chemistry. Furthermore, the molecular specificity of the surface enhancement process is not well understood. Therefore, a truly label free method for monitoring and characterising the cellular uptake and subcellular localisation of nanoparticles in general is still required.

It has been demonstrated that, based on the spectroscopic fingerprint of polystyrene, Raman spectroscopy, coupled with routine multivariate analyses such as

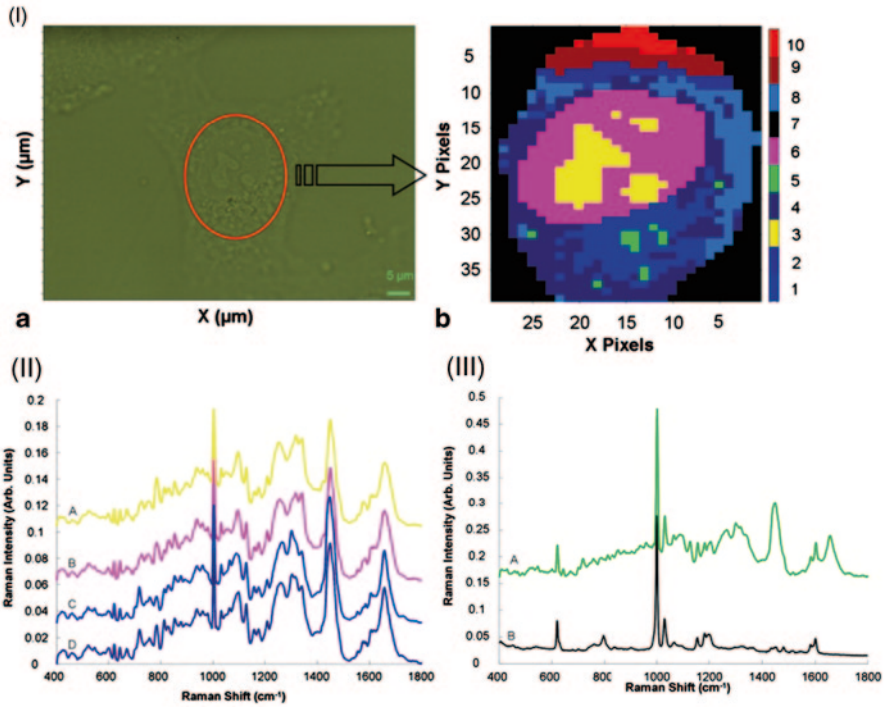


Fig. 13.19 I **a** Microscopic image of an A549 cell, showing the reduced area identified for spectral mapping. **b** K-means cluster map of the Raman profile of the same reduced area. II K-means spectra of clusters 3 (*A* representing nucleoli), 6 (*B* representing nucleus), 1 and 4 (*C* and *D*, both from the cytoplasm). Spectra are offset for clarity. III K-means spectrum of Cluster 5 (*A*), compared to the Raman spectrum polystyrene nanoparticles (*B*). Spectra are offset for clarity. (Reproduced from [151])

K-means clustering, is a more specific probe to detect and localise nanoparticles within cells [151]. Furthermore, Principal Component Analysis can be employed to differentiate the local environment of the nanoparticles from the cytoplasm. For convenience, formalin fixed (A549) cells were employed. Raman mapping was performed in immersion using a x100 Olympus LUMPlanF1 objective, a wavelength of 785 nm, and a step size of 0.75 μm. As seen in Fig. 13.13, using a 100x objective, the visible image obtained from the cells is rather detailed (Fig. 13.19 Ia). The nucleus is clearly identifiable and the nucleoli can also be seen. The membrane of the cells is well defined and the cytoplasm presents different structures and the region containing the endoplasmic reticulum can be easily recognised. As in the previous section, K-means cluster analysis of the Raman spectral map using 10 clusters clearly differentiates the sub-cellular regions of interest (Fig. 13.19 Ib).

The nucleoli spectra form a specific group (cluster 3), distinct from the nucleus (cluster 6). The corresponding mean spectra are displayed in Fig. 13.19 II. The cytoplasm is represented by a number of different clusters (1, 2, 4, 8, 9, 10). Notably,

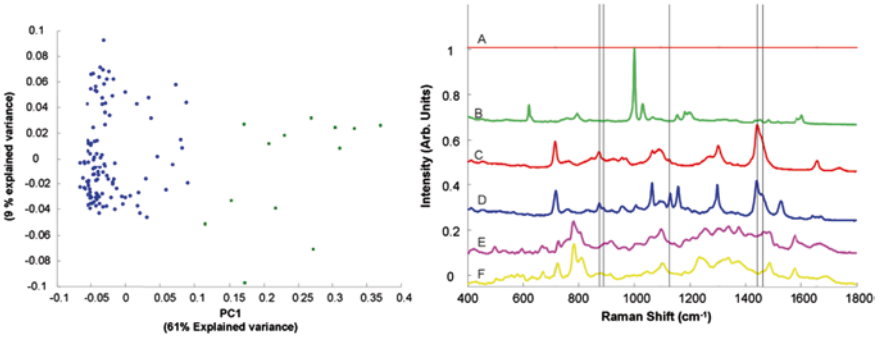


Fig. 13.20 *Top*; Scatter plot of the PCA of spectra associated with clusters 1 (*Blue*) and 5 (*Green*). *Bottom*; (*A*) Loading of PC1 (*B*) Raman Spectrum of polystyrene nanoparticles (*C*) of L-α-Phosphatidylcholine, (*D*) of Sphingomyelin, (*E*) of DNA and (*F*) of RNA. (Reproduced from [151])

nanopolystyrene can now be localised in the cytoplasm according to cluster 5, the corresponding mean spectrum of which is presented in Fig. 13.19 IIIa. Although some cellular features clearly contribute to the cluster spectrum, it is dominated by specific features originating from polystyrene (Fig. 13.19 IIIb). K-means cluster analysis yields the average spectra for the identified clusters. These average spectra can contain contributions from points on the boundaries between the cellular regions and a direct comparison does not always easily facilitate identification of the differences in the biochemical composition. Cluster 5 clearly identifies the presence of polystyrene nanoparticles localised in the cytoplasm, but the averaged spectrum also exhibits clear contributions of the biological environment. PCA can be employed to highlight the biochemical differences between the subcellular regions [150]. In Fig. 13.19 Ib, cluster 5 is predominantly surrounded by cytoplasmic regions grouped within cluster 1. Figure 13.20 (top) is a scatter plot of the PCA of spectra associated with K-means clusters 1 and 5. For visualisation purposes, spectra associated with cluster 1 are coloured blue, while those from cluster 5 are coloured green, and although the data is rather dispersed, some differentiation between the two clusters according to PC1 is apparent. PC1 accounts for 61 % of the variance and, as shown in the loadings plot of Fig. 13.20 (bottom), it is dominated by positive contributions of polystyrene (Fig. 13.20b). Notably, the negative contributions of PC1 match well with the Raman features of lipids such as L-α-Phosphatidylcholine (Fig. 13.20c) and also sphingomyelin (Fig. 13.20d). For comparison, the Raman spectra of DNA (Fig. 13.20e) and RNA (Fig. 13.20f) are also shown, and it is clear that they exhibit few or no similarities with PC1. High loadings of PC1 for cluster 5 are therefore an indication of the dominance of polystyrene. Positive loadings of PC1 for some of Cluster 1 may indicate regions of overlap between the localised nanoparticles and the surrounding cytoplasm, which, although they have significant polystyrene contributions, are represented in the K-means cluster analysis by the average spectrum of cluster 1.

An established mechanism for the transport of nanoparticles in cells is via endosomes and later lysosomes [170]. A number of studies have also demonstrated that, in the later stages of trafficking, individually endocytosed nanoparticles accumulate in larger multivescular bodies [165] and they have also been seen to be localized in the endoplasmic reticulum and Golgi apparatus [171]. Notably, in Fig. 13.19 Ia, the majority of the nanoparticles, identified in green as cluster 5, are surrounded by cytoplasmic regions of cluster 1, which is relatively spatially extensive, particularly compared to typical sizes of lysosomes (1–2 μm). In the white light image of Fig. 13.19 Ib, this region can be identified as the Endoplasmic Reticulum (ER), which may be expected to be rich in lipidic compounds, and a PCA of the different regions of the cytoplasm confirms this. Cluster 4, on the other hand is relatively rich in nucleic acids, which may be due to the presence of mitochondria. Co-staining the cells to highlight the ER and examination using confocal fluorescence microscopy indicates a strong correlation between the ER stain and the fluorescence of the nanoparticle label, and any locally concentrated fluorescence which may be indicative of encapsulated nanoparticles is contained within the ER [151]. The analysis supports the attribution of the subcellular environment of K-means cluster 1 as endoplasmic reticulum.

The study represents an initial proof of concept, to demonstrate the potential use of Raman spectroscopy to identify and locate nanoparticles in cells and to probe their subcellular environment. No attempt has, as yet, been made to quantify the amount of nanoparticles located or the sensitivity and detection limit of the technique. Instrumentally, the spatial resolution can be improved by moving to lower wavelength sources, although the excitation of any fluorescence from the nanoparticles and the cell constituents should be avoided. However, as it scales linearly with wavelength, a move from 785 to 473 nm should result in an improvement in lateral resolution by a factor of ~1.7. Ultimately, the confocal mode of operation can provide 3D localisation of the nanoparticles at submicron resolution.

Other multivariate analytical approaches have also been applied in the field of Raman microspectroscopy of cells. In addition to KMCA, other clustering methods such as Fuzzy C means clustering (FCM) and hierarchal cluster analysis (HCA) have been used to separate the cellular Raman data into clusters and subsequently reshape the data into images [172, 173]. However, as highlighted by Headegaad et al., these approaches have their own limitations. In particular, boundaries between sub-cellular features can often result in the addition of extra clusters with mixed spectral signatures. This can be overcome by increasing the number of clusters; however, this in turn can result in added complexity to interpretation and inaccuracies in regional separation. Additionally, the reproducibility of these methods can also be questioned as the starting point for the centroid based KMCA and FCM is subjective [172].

PCA and vertex component analysis (VCA) have also been used to separate out distinct regions of the cell. With regards to PCA, separation is based on the variances between the spectra in the data set, the majority of the variance being described by the first three principal components [172]. Thus, the score values can be used to construct a composite image of the cell in which the biochemical contributions of

each component are described by the corresponding loadings plot. Unlike KMCA and FCM, PCA identifies quite accurately the boundaries between each feature. However, the images generated suffer from inferior contrast and in some instances interpretation may be difficult as biochemical features may be spread across different loadings.

VCA is another method which has been used for similar analytical purposes. In brief, VCA computes a linear combination of supposed pure component spectra which are termed endmember spectra. As described in Miljkovic et al., the endmember spectra are acquired under the assumption that the most extreme data points in the dataset are representative of pure component spectra [173]. However, it has been pointed out that the endmembers generated are not truly representative of the pure component they describe in the data set and can often contain a mixture of biochemical constituents i.e. DNA and proteins. While this is representative of the true nature of nucleic acids *in-situ*, it could lead to inaccuracies in interpretation.

A more recent study introduced the technique of spectral cross correlation to screen spectral maps for the spatial distribution of specific spectral signatures, including those of polystyrene nanoparticles, RNA and lipids, and argued that it was more effective than Classical least squared analysis, both supervised and unsupervised, which is commonly available on instrument software [174].

As in the case of multivariate techniques for disease diagnostics, there is a myriad of potential techniques to explore the spatial variations of the spectral profiles with heterogeneous structures such as biological cells. A truly quantitative comparison of the different techniques is lacking, however.

13.7 Conclusions

Since their origins in the nineteenth and twentieth centuries, the techniques of Infrared absorption and Raman scattering spectroscopy have come a long way. Technological developments have rendered them routine benchtop techniques for chemical and process analysis. Although their applications for the analysis of complex mixtures have raised some challenges, an understanding of the fundamental physical processes and interactions has helped the scientific community address these challenges. In biomedical applications, the chemical and physical inhomogeneities of cells and tissues have presented many such challenges, not easily understood by a multidisciplinary community. However, concerted efforts through collaborative programmes such as the EU FP6 Special Support Action DASIM [175] have helped understand the origins of spectral artefacts which can mask or obscure the true information contained within a spectral profile, such that such effects can be minimised by improved sample presentation and/or data preprocessing.

Over the past decades, the potential of vibrational spectroscopy for biomedical applications has been extensively explored and demonstrated. The prospect is of a truly label free, non-invasive, non destructive, rapid, automated and objective technology for disease diagnostics and beyond. Disease diagnostics can be made on

the basis of true changes in biochemistry rather than for example cellular and tissue morphology, and disease origin and progression can be better understood. Gerwin Puppels, one of the pioneers of the field and founder of River Diagnostics likens the developing technology to that of Dr. McCoy of Star Trek, whereby a simple optical probe can scan a body for all ailments. While it may seem improbable, the emergence of spectroscopic probes for *in vivo* analyses and the development of applications for the analysis of bodily fluids and identification of spectral markers of disease may mean the realisation of real clinical applications in the near future.

Beyond disease diagnostics, EU legislation is placing increasing emphasis on *in vitro* screening techniques for the replacement, reduction and refinement of animal experimentation in scientific studies. High Content Analysis techniques for automation and higher throughput of screening based on chemical and immuno assays have developed significantly over the past decade. However, the assays are highly specific and the process of assaying is extremely expensive. In the realm of *in vitro* diagnostics and therapeutics, vibrational spectroscopy presents as realistic alternative.

The development of applications of vibrational spectroscopy for biomedical/chemical applications has relied heavily on the development and application of multivariate statistical and data mining techniques. A single spot on a cell or tissue contains a multitude of different biochemical components in itself, and the composition and morphology of this complex melange can vary dramatically over a length scale which is comparable to the sampled spot size. Cell to cell, tissue to tissue, or patient to patient variability must also be considered. The challenge of extracting relevant information is therefore immense, and a myriad of techniques has been applied and adapted as exemplified in a multitude of techniques. Given the impact of the conclusions derived, particularly in medical diagnostics, it is imperative that the combination of spectral analysis and data processing yield the "right answer". In a real life scenario, however, it is impossible to know for certain what the "right answer" is however. In this context, complex simulated data sets which incorporate instrumental variabilities and measurement artefacts, but also sample to sample variabilities may play a critical role in validating and optimising the data analytical techniques. With a combination of optimised spectroscopic technologies, sample presentation protocols and data processing and analysis techniques, the future for vibrational spectroscopy is bright, for disease diagnostics and beyond.

References

1. http://coolcosmos.ipac.caltech.edu/cosmic_classroom/classroom_activities/herschel_bio.html
2. Rieke GH (2009) History of infrared telescopes and astronomy. Exp Astron 25:125
3. Smith AL (1979) Applied Infrared spectroscopy: fundamentals, techniques, and analytical problemsolving. Wiley, New York
4. Mertz L (1965) Astron J 70:548
5. Griffiths PR, Curbelo R, Foskett CT, Dunn ST (1970) Analytical instrumentation (Inst. Society of America), 8, II4 (1970)
6. Messerchmidt R, Harthcock M (eds) (1988) Infrared microscopy, theory and applications. Marcel Dekker, New York

7. http://www.mtholyoke.edu/~mlyount/MySites/ForensicSpectroscopy/ForensicApps.html
8. USA Today, June 23rd 1993, pg8A
9. Lyon RC, Jefferson EH, Ellison CD, Buhse LF, Spencer JA, Nasr MM, Hussain AS Exploring Pharmaceutical Applications of NearInfrared Technology. American Pharmaceutical Review, http://americanpharmaceuticalreview.com/ViewArticle.aspx?ContentID=225
10. Raman CV, Krishnan KS (1928) Nature 121:501
11. Adjouri C, Elliasmine A, Le Duff Y (1996) Spectroscopy 44:46
12. Puppels GJ, Olminkhof JHF, Segers-Nolten GMJ, Otto C, de Mul FFM, Greve J (1991) Exp Cell Res 195:361
13. Jackson M, Mantsch H (1995) CRC Crit Rev Biochem Mol Biol 30:95
14. Casal HL, Mantsch HH (1984) Biochim Biophys Acta 779:381
15. Mathlouthi M, Koenig JL (1986) Adv Carbohydr Chem Biochem 44:7
16. Taillandier E, Liquier J, Taboury JA (1985) In: Clark RJH, Hester RE (eds), Advances in spectroscopy: advances in infrared and Raman spectroscopy, vol 12. Wiley, New York, p 65
17. Wong P, Lacele S, Yadzi H (1993) Appl Spectroscopy 47:1830
18. Dukor R (2002) Vibrational spectroscopy in the detection of cancer. In: Chalmers JM, Griffiths PR (eds) Handbook of vibrational spectroscopy, vol 5. Wiley, Chichester
19. Gazi E, Baker M, Dwyer J, Lockyer NP, Gardner P, Shanks JH, Reeve RS, Hart C, Clarke NW, Brown M (2006) Eur Urol 50:750–761
20. Fernandez DC, Bhargava R, Hewitt SM, Levin IW (2005) Nature Biotechnol 23:469
21. Lord RC, Yu NT (1970) J Mol Biol 20:50924
22. Tobin MC (1998) Science 161:68
23. Walton AG, Deveney MJ, Koenig JL (1970) Calcified Tissue Int 6:162
24. Yu NT, Jo BH, Chang RCC, Huber JD (1974) Arch Biochem Biophys 160:614
25. Puppels GJ, Breve J (1996) Biomedical applications of spectroscopy. In: Clark RHJ, Hester RE (eds) Advances in spectroscopy, vol 25. Wiley, New York
26. Gniadecka M, Wulf HC, Nielsen OF, Christensen DH, Hercogova J (1997) Photochem Photobiol 66:418
27. Krishna CM, Sockalingum GD, Venteo L, Bhat RA, Kustagi P, Pluot M, Manfait M (2002) Biopolymers 5:269
28. Smith J, Kendall C, Sammon A, ChristieBrown J, Stone N (2003) Technol Cancer Res Treatment 2:327
29. Molckovsky A, Song LMWK, Shim MG, Marcon NE, Wilson BC (2003) Gastrointest Endosc 57:396
30. Hanlon EB, Manoharan R, Koo TW, Shafer KE, Motz JT, Fitzmaurice M, Kramer JR, Itzkan I, Dasari RR, Feld MS (2000) Phys Med Biol 45:R1–R59
31. Caspers PJ, Lucassen GW, Wolthuis R, Bruining HA, Puppels GJ (1999) Biospectroscopy 4:S31–39
32. Utzinger U, MahadevanJansen A, Hinzelman D, Follen M, Richards-Kortum R (2001) Appl Spectrosc 55:955
33. http://www.chem.neu.edu/web/faculty/Prof.MaxDiem.html
34. Lyng FM, Faoláin EÓ, Conroy J, Meade A, Knief P, Duffy B, Hunter M, Byrne J, Kelehan P, Byrne HJ (2007) Exp Mol Pathol 82:121
35. Parkin DM, Bray F, Ferlay J, Pisani P (2005) Global cancer statistics 2002. CA Cancer J Clin 55:74–108
36. Nanda K, McCrory DC, Myers ER, Bastian LA, Hasselblad V, Hickey JD, Matchar DB (2000) Ann Intern Med 132:810
37. Willis BH, Barton P, Pearmain P, Bryan S, Hyde C (2005) Health Technol Assess 9:1–207
38. Bolger N, Heffron C, Regan I, Sweeney M, Kinsella S, McKeown M, Creighton G, Russell J, O'Leary J (2006) Acta Cytol 50:483
39. Koliopoulos G, Arbyn M, Martin-Hirsch P, Kyrgiou M, Prendiville W, Paraskevaidis E (2006) Gynecol Oncol 104:232
40. Chang SK, Pavlova T, Marin N, Follen M, Richards-Kortum R (2005) Gynecol Oncol 99:S61–63

41. Gurjar RS, Backman V, Perelman LT, Georgakoudi I, Badizadegan K, Itzkan I, Dasari RR, Feld MS (2001) Nat Med 7:1245
42. Escobar PF, Belinson JL, White A, Shakhova NM, Feldchtein FI, Kareta MV, Gladkova ND (2004) Int J Gynecol Cancer 14:470
43. Zuluaga AF, Follen M, Boiko I, Malpica A, Richards-Kortum R (2005) Am J Obstet Gynecol 193:83
44. Carlson K, Pavlova I, Collier T, Descour M, Follen M, Richards-Kortum R (2005) Gynecol Oncol 99:S84
45. Wood BR, Chiriboga L, Yee H, Quinn MA, McNaughton D, Diem M (2003) Gynecol Oncol 93:59
46. Mordechai S, Sahu RK, Hammody Z, Mark S, Kantarovich K, Guterman H, Podshyvalov A, Goldstein J, Argov S (2004) J Microscopy-Oxford 215:86
47. Romeo M, Matthaus C, Miljkovic M, Diem M (2004) Biopolymers 74:168
48. Mahadevan-Jansen A, Ramanujam N, Malpica A, Thomsen S, Utzinger U, Richards-Kortum R (1998) Photochem Photobiol 68:123
49. Faoláin EÓ, Hunter MB, Byrne JM, Kelehan P, Lambkin HA, Byrne HJ, Lyng FM (2005) J Histochem Cytochem 53:121
50. Jackson JE (1991) A user's guide to principal components. Wiley, New York
51. Otto M (1999) Chemometrics: statistics and computer applications in analytical chemistry. Wiley, New York
52. Crow P, Molckovsky A, Stone N, Uff J, Wilson B, Wong-KeeSong LM (2005) Urology 65:1126
53. Haka AS, Shafer-Peltier KE, Fitzmaurice M, Crowe J, Dasari RR, Feld MS (2002) Cancer Res 62:5375
54. Nijssen A, Bakker Schut TC, Heule F, Caspers PJ, Hayes DP, Neumann MH, Puppels GJ (2002) J Invest Dermatol 119:64
55. UK Patent Application No. GB0823071.6 (2008) A diagnostic instrument for cervical cancer screening. (December 2008)
56. International Patent, Lyng FM, Byrne HJ, Faoláin EÓ, Clarke C, Ostrowska KM (2010) A diagnostic instrument for cervical cancer screening. Patent no WO2010070133
57. Almond ML, Hutchings J, Kendall C, Stone N, Barr H (2011) Brit J Surg 98:35
58. Bassan P, Byrne HJ, Bonnier F, Lee J, Dumas P, Gardner P (2009) Analyst 134:1586
59. Bassan P, Byrne HJ, Lee J, Bonnier F, Clarke C, Dumas P, Gazi E, Brown MD, Clarke NW, Gardner P (2009) Analyst 134:1171
60. Bassan P, Kohler A, Martens H, Lee J, Byrne HJ, Dumas P, Gazi E, Brown M, Clarke N, Gardner P (2010) Analyst 135:268
61. Bonnier F, Mehmood A, Knief P, Meade A, Hornebeck W, Lambkin H, Flynn K, McDonagh V, Healy C, Lee TC, Lyng FM, Byrne HJ (2011) J Raman Spectrosc 42:888
62. Bonnier F, Ali SM, Knief P, Lambkin H, Flynn K, McDonagh V, Healy C, Lee TC, Lyng FM, Byrne HJ (2012) Vib Spectrosc 61:124
63. Ostrowska KM, Malkin A, Meade AD, O'Leary J, Martin C, Spillane C, Byrne HJ, Lyng FM (2010) Analyst 135:3087
64. Ostrowska KM, Garcia A, Meade AD, Malkin A, Okewumi I, O'Leary JJ, Martin C, Byrne HJ, Lyng FM (2011) Analyst 136:1365
65. Cuschieri KS, Cubie HA, Whitley MW, Gilkison G, Arends MJ, Graham C, McGoogan E (2005) J Clin Pathol 58:946
66. Tjalma WA, Van Waes TR, Van den Eeden LE, Bogers JJ (2005) Best Pract Res Clin Obstet Gynaecol 19:469
67. Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, Snijders PJ, Peto J, Meijer CJ, Munoz N (1999) J Pathol 189:12
68. Mayrand MH, Duarte-Franco E, Rodrigues I, Walter SD, Hanley J, Ferenczy A, Ratnam S, Coutlee F, Franco EL (2007) N Engl J Med 357:1579
69. Jess PR, Smith DD, Mazilu M, Dholakia K, Riches AC, Herrington CS (2007) Int J Cancer 121:2723

70. Kelly JG, Cheung KT, Martin C, O'Leary JJ, Prendiville W, Martin-Hirsch PL, Martin FL (2010) Clin Chim Acta 411:1027
71. Murphy N, Ring M, Heffron CC, King B, Killalea AG, Hughes C, Martin CM, McGuinness E, Sheils O, O'Leary JJ (2005) J Clin Pathol 58:525
72. Max JJ, Chapados C (2009) J Chem Phys 131:184505
73. Martin CM, Kehoe L, Spillane CO, O'Leary JJ (2007) Mol Diagn Ther 11:277
74. Kalof A, Evans M, Simmons-Arnold L, Beatty B, Cooper K (2005) Am J Surg Pathol 29:674
75. Dray M, Russell P, Dalrymple C, Wallman N, Angus G, Leong A, Carter J, Cheerala B (2005) Pathology 37:112
76. Yu T, Ferber MJ, Cheung TH, Chung TK, Wong YF, Smith DI (2005) Cancer Genet Cytogenet 158:27
77. Ozgul N, Cil A, Bozdayi G, Usubutun A, Bulbul D, Rota S, Kose M, Biri A, Haberal A (2008) J Obstet Gynaecol Res 34:865
78. Murphy N, Ring M, Heffron C, King B, Killalea A, Hughes C, Martin C, McGuinness E, Sheils O, O'Leary J (2005) J Clin Pathol 58:525
79. Missaoui N, Hmissa S, Frappart L, Trabelsi A, Ben Abdelkader A, Traore C, Mokni M, Yaacoubi MT, Korbi S (2006) Virchows Arch 448:597
80. Murphy N, Ring M, Killalea AG, Uhlmann V, O'Donovan M, Mulcahy F, Turner M, McGuinness E, Griffin M, Martin C, Sheils O, O'Leary JJ (2003) J Clin Pathol 56:56
81. Klaes R, Friedrich T, Spitkovsky D, Ridder R, Rudy W, Petry U, Dallenbach-Hellweg G, Schmidt D, von Knebel Doeberitz M (2001) Int J Cancer 92:276
82. Agoff SN, Lin P, Morihara J, Mao C, Kiviat NB, Koutsky LA (2003) Mod Pathol 16:665
83. Murphy N, Heffron CC, King B, Ganuguapati UG, Ring M, McGuinness E, Sheils O, O'Leary JJ (2005) Virchows Arch 445:610
84. Murphy N, Ring M, Heffron CC, King B, Killalea AG, Hughes C, Martin CM, McGuinness E, Sheils O, O'Leary JJ (2005) J Clin Pathol 58:525
85. Wang SS, Trunk M, Schiffman M, Herrero R, Sherman ME, Burk RD, Hildesheim A, Bratti MC, Wright T, Rodriguez AC, Chen S, Reichert A, von Knebel Doeberitz C, Ridder R, von Knebel Doeberitz M (2004) Cancer Epidemiol Biomarkers Prev 13:1355
86. Movasaghi Z, Rehman S, Rehman IU (2007) Appl Spectrosc Rev 42:493
87. Black JW, Leff P (1983) Proc R Soc Lond B Biol Sci 220:141
88. Lizard G, Chignol MC, Chardonnet Y, Souchier C, Bordes M, Schmitt D, Revillard JP (1993) J Immunol Methods 157:31
89. Mehés G, Speich N, Bollmann M, Bollmann R (2004) Pathol Oncol Res 10:142
90. Burd EM (2003) Clin Microbiol Rev 16:1
91. Neviliappan S, Fang Kan L, Tiang Lee Walter T, Arulkumaran S, Wong PT (2002) Gynecol Oncol 85:170
92. Yazdi HM, Bertrand MA, Wong PT (1996) Acta Cytol 40:664
93. Suprynowicz FA, Disbrow GL, Krawczyk E, Simic V, Lantzky K, Schlegel R (2008) Oncogene 27:1071
94. Wong PT, Wong RK, Caputo TA, Godwin TA, Rigas B (1991) Proc Natl Acad Sci USA 88:10988
95. Martens H, Martens M (2000) Food Qual Prefer 11:5
96. Westad F, Martens H (2000) J Near Infrared Spectrosc 8:117
97. Doorbar J (2006) Clin Sci (Lond) 110:525
98. Nawaz H, Bonnier F, Knief P, Howe O, Lyng FM, Meade AD, Byrne HJ (2010) Analyst 135:3070
99. Nawaz H, Bonnier F, Meade AD, Lyng FM, Byrne HJ (2011) Analyst 136:2450
100. Owen CA, Selvakumaran J, Notingher I, Jell G, Hench LL, Stevens MM (2006) J Cell Biochem 99:178
101. Draux F, Jeannesson P, Beljebbar A, Tfayli A, Fourre N, Manfait M, Sule-Suso J, Sockalingum GD (2009) Analyst 134:542
102. Ling J, Weitman SD, Miller MA, Moore RV, Bovik AC (2002) Appl Opt 41:6006

103. Ling JY, Yang QZ, Luo SS, Li Y, Zhang CK (2005) Chin Chem Lett 16:71
104. Notingher I, Green C, Dyer C, Perkins E, Hopkins N, Lindsay C, Hench LL (2004) J R Soc Interface 1:79
105. Verrier S, Notingher I, Polak JM, Hench LL (2004) Biopolymers 157
106. Yang Y, Sule-Suso J, Sockalingum GD, Kegelaer G, Manfait M, El Haj AJ (2005) Biopolymers 78:311
107. Kosmider B, Wojcik I, Osiecka R, Bartkowiak J, Zyner E, Ochocki J, Liberski P (2005) Invest New Drugs 23:287
108. Vrana O, Masek V, Drazan V, Brabec V (2007) J Struct Biol 159:1
109. Ulukaya E, Ozdikicioglu F, Oral AY, Demirci M (2008) Toxicol In Vitro 22:232
110. Cordes N, Beinke C, Plasswilm L, van Beuningen D (2004) Strahlenther Onkol 180:157
111. Baik MH, Friesner RA, Lippard SJ (2003) J Am Chem Soc 125:14082
112. Dunham SU, Turner CJ, Lippard SJ (1998) J Am Chem Soc B:5395
113. Gelasco A, Lippard SJ (1998) Biochemistry 37:9230
114. Takahara PM, Rosenzweig AC, Frederick CA, Lippard SJ (1995) Nature 377:649
115. Meade AD, Clarke C, Byrne HJ, Lyng FM (2010) Radiat Res 173:225
116. Varmuza K, Filzmoser P (2009) Introduction to multivariate statistical analysis in chemometrics. CRC Press, Taylor and Francis Group, Boca Raton
117. Krishna CM, Kegelaer G, Adt I, Rubin S, Kartha VB, Manfait M, Sockalingum GD (2006) Biopolymers 82:462
118. Le Gal JM, Morjani H, Manfait M (1993) Cancer Res 53:3681
119. Bonnier F, Meade AD, Merzha S, Knief P, Bhattacharya K, Lyng FM, Byrne HJ (2010) Analyst 135:1697
120. Fu Y, Wang H, Shi R, Cheng J-X (2007) Biophys J 92:3251
121. Colella G, Marchini S, D'Incalci MD, Brown R, Broggini M (1999) Br Cancer J 80:338
122. Bonnier F, Knief P, Lim B, Meade AD, Dorney J, Bhattacharya K, Lyng FM, Byrne HJ (2010) Analyst 135:3169
123. Notingher I, Verrier S, Haque S, Polak JM, Hench LL (2003) Biopolymers 72:230
124. Notingher I, Hench LL (2006) Expert Rev Med Devices 3:215
125. Puppels GJ, de Mul FF, Otto C, Greve J, Robert-Nicoud M, Arndt-Jovin DJ, Jovin TM (1990) Nature 347:301
126. Puppels GJ, Garritsen HS, Segers-Nolten GM, de Mul FF, Greve J (1991) Biophys J 60:1046
127. Miljkovic M, Chernenko T, Romeo MJ, Bird B, Matthaus C, Diem M (2010) Analyst 135:2002
128. Uzunbajakava N, Lenferink A, Kraan Y, Willekens B, Vrensen G, Greve J, Otto C (2003) Biopolymers 72:1
129. Draux F, Gobinet C, Sule-Suso J, Trussardi A, Manfait M, Jeannesson P, Sockalingum GD (2010) Anal Bioanal Chem 397:2727
130. Meade AD, Clarke C, Draux F, Sockalingum GD, Manfait M, Lyng FM, Byrne H (2010) J Anal Bioanal Chem 396:1781
131. Weaver VM, Howlett AR, Langton-Webster B, Petersen OW, Bissell MJ (1995) Semin Cancer Biol 6:175
132. Petersen OW, Ronnov-Jessen L, Howlett AR, Bissell MJ (1992) Proc Natl Acad Sci USA 89:9064
133. Prestwich GD (2008) Acc Chem Res 41:139
134. Smalley KS, Lioni M, Herlyn M (2006) In Vitro Cell Dev Biol: Anim 42:242
135. Curtis A, Wilkinson C (1997) Biomaterials 18:1573
136. Giannakakou P, Nakano M, Nicolaou KC, O'Brate A, Yu J, Blagosklonny MV, Greber UF, Fojo T (2002) Proc Natl Acad Sci USA 99:10855
137. Kotorashvili A, Russo SJ, Mulugeta S, Guttentag S, Beers MF (2009) J Biol Chem 284:16667
138. Hsia TC, Yang JS, Chen GW, Chiu TH, Lu HF, Yang MD, Yu FS, Liu KC, Lai KC, Lin CC, Chung JG (2009) Anticancer res 29:309

139. Pettus BJ, Bielawska A, Subramanian P, Wijesinghe DS, Maceyka M, Leslie CC, Evans JH, Freiberg J, Roddy P, Hannun YA, Chalfant CE (2004) J Biol Chem 279:11320
140. Grewal S, Ponnambalam S, Walker JH (2003) J Cell Sci 116:2303
141. MacQueen JB (1967) Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability
142. Koljenovic S, Bakker Schut TC, van Meerbeeck JP, Maat AP, Burgers SA, Zondervan PE, Kros JM, Puppels GJ (2004) J Biomed Opt 9:1187
143. Lasch P, Haensch W, Naumann D, Diem M (2004) Biochimica et Biophysica Acta 1688:176
144. Ly E, Piot O, Durlach A, Bernard P, Manfait M (2009) Analyst 134:1208
145. Hedegaard M, Krafft C, Ditzel HJ, Johansen LE, Hassing S, Popp J (2010) Anal Chem 82:2797
146. Taleb A, Diamond J, McGarvey JJ, Beattie JR, Toland C, Hamilton PW (2006) J Phys Chem 110:19625
147. Short KW, Carpenter S, Freyer JP, Mourant JR (2005) Biophys J 88:4274
148. Swain RJ, Jell G, Stevens MM (2008) J Cell Biochem 104:1427
149. Krishna CM, Sockalingum GD, Venteo L, Bhat RA, Kushtagi P, Pluot M, Manfait M (2005) Biopolymers 79:269
150. Bonnier F, Byrne HJ (2012) Analyst 137:322
151. Dorney J, Bonnier F, Garcia A, Casey A, Chambers G, Byrne HJ (2012) Analyst 137:1111
152. Berntsen P, Park CY, Rothen-Rutishauser B, Tsuda A, Sager TM, Molina RM, Donaghey TC, Alencar AM, Kasahara DI, Ericsson T, Millet EJ, Swenson J, Tschumperlin DJ, Butler JP, Brain JD, Fredberg JJ, Gehr P, Zhou EH (2010) J R Soc Interface 7:S331
153. Lundqvist M, Stigler J, Elia G, Lynch I, Cedervall T, Dawson KA (2008) Proc Natl Acad Sci U.S.A. 105:14265
154. Clift MJD, Rothen-Rutishauser B, Brown DM, Duffin R, Donaldson K, Proudfoot L, Guy K, Stone V (2008) Toxicol Appl Pharmacol 232:418
155. Rejman J, Oberle V, Zuhorn IS, Hoekstra D (2004) Biochem J 377:159
156. Rothen-Rutishauser B, Muhlfeld C, Blank F, Musso C, Gehr P (2007) Part Fibre Toxicol 4:9
157. Naha PC, Bhattacharya K, Tenuta T, Dawson KA, Lynch I, Garcia A, Lyng FM, Byrne HJ (2010) Toxicol Lett 198:134
157. Roy I, Ohulchanskyy TY, Bharali DJ, Pudavar HE, Mistretta RA, Kaur N. Prasad PN (2005) Proc Natl Acad Sci USA 102:279
158. Cang H, Xu CS, Montiel D, Yang H (2007) Opt Lett 32:2729
159. Salvati A, Aberg C, Dos Santos T, Varela J, Pinto P, Lynch I, Dawson KA (2011) Nanomedicine 7:818
160. Suh H, Jeong B, Liu F, Kim SW (1998) Pharm Res 15:1495
161. Yin Win K, Feng S-S (2005) Biomaterials 26:2713
162. Davoren M, Herzog E, Casey A, Cottineau B, Chambers G, Byrne HJ, Lyng FM (2007) Toxicol in Vitro 21:438
163. Shapero K, Fenaroli F, Lynch I, Cottell DC, Salvati A, Dawson KA (2011) Mol BioSyst 7:371
164. Knief P, Clarke C, Herzog E, Davoren M, Lyng FM, Meade AD, Byrne HJ (2009) Analyst 134:1182
165. Kneipp K, Haka AS, Kneipp H, Badizadegan K, Yoshizawa N, Boone C, Shafer-Peltier KE, Motz JT, Dasari RR, Feld MS (2002) Appl Spectrosc 56:150
166. Kneipp J, Kneipp H, McLaughlin M, Brown D, Kneipp K (2006) Nano Lett 6:2225
167. Kneipp J, Kneipp H, Rajadurai A, Redmond RW, Kneipp K (2009) J Raman Spectrosc 40:1
168. Kneipp J, Kneipp H, Rice WL, Kneipp K (2005) Anal Chem 77:2381
169. Kneipp J, Kneipp H, Wittig B, Kneipp K (2010) J Phys Chem C 114:7421
170. Nel AE, Madler L, Velegol D, Xia T, Hoek EMV, Somasundaran P, Klaessig F, Castranova V, Thompson M (2009) Nat Mater 8:543

171. Chang MY, Shiau AL, Chen YH, Chang CJ, Chen HH, Wu CL (2008) Cancer Sci 99:1479
172. Hedegaard M, Matthaus C, Hassing S, Krafft C, Diem M. Popp J (2011) Theor Chem Acc 130:1249
173. Miljković M, Chernenko T, Romeo MJ, Bird B, Matthäus C, Diem M (2010) Analyst 135:2002
174. Keating ME, Bonnier F, Byrne HJ, (2012) Analyst DOI: 10.1039/c2an36169h
175. http://www.dasim.eu/website.php
176. Benevolo M, Mottolese M, Marandino F, Vocaturo G, Sindico R, Piperno G, Mariani L, Sperduti I, Canalini P, Donnorso RP, Vocaturo A (2006) Immunohistochemical expression of p16(INK4a) is predictive of HR-HPV infection in cervical low-grade lesions. Mos Pathol 19:384–391
177. Passamonti B, Gustinucci D, Recchia P, Bulletti S, Carlani A, Cesarini E, D'Amico MR, D'Angelo V, Di Dato E, Martinelli N, Malaspina M, Spita N (2010) Expression of p16 in abnormal pap-tests as an indicator of CIN2+ lesions: a possible role in the low grade ASC/US and L/SIL (Ig) cytologic lesions for screening prevention of uterine cervical tumours. Pathologica 102:6–11
178. Negri G, Bellisano G, Zannoni GF, Rivasi F, Kasal A, Vittadello F, Antoniazzi S, Faa G, Ambu R, Egarter-Vigl E (2008) p16 ink4a and HPV L1 immunohistochemistry is helpful for estimating the behavior of low-grade dysplastic lesions of the cervix uteri. Am J Surg Pathol 32:1715–1720
179. Focchi GR, Silva ID, Nogueira-de-Souza NC, Dobo C, Oshima CT, Stavale JN (2007) Immunohistochemical expression of p16(INK4A) in normal uterine cervix, nonneoplastic epithelial lesions, and low-grade squamous intraepithelial lesions. J Low Genit Tract Dis 11:98–104

Chapter 14
Quantitative Analysis of Disease Biomarkers Using Surface-Enhanced Raman Scattering Spectroscopy

Hyangah Chon and Jaebum Choo

Abstract In this chapter, we describe the SERS-based immunoassay technique for the highly sensitive detection of various disease biomarkers. The fundamentals of the technique are briefly presented, and three different SERS assay platforms and their biomedical applications are described. The novel SERS-based assay technique is expected to be a powerful clinical tool for early disease diagnosis.

Keywords Surface-enhanced Raman scattering (SERS) • Immunoassay • Biomarker • Early diagnosis • Microfluidics

14.1 Introduction

Immunoassays are a fast and cost-effective method that can be applied to biochemical studies, clinical diagnostics, food safety tests and environmental monitoring [1–3]. Enzyme-linked immunosorbent assay (ELISA) is the most widely used screening tool for validating candidate protein markers [4]. This technique allows for the simultaneous estimation of the concentration of numerous biomarkers in a small volume. ELISAs also have several significant advantages over other conventional assay methods, such as a high-throughput performance, cost-effectiveness and facile clinical applications. Nonetheless, this assay technique has some technical limitations. Fluorescence detection using molecular labels is commonly used to identify a specific antigen-antibody interaction in an immunoassay; however, this has several drawbacks including a poor limit of detection, photobleaching, and limited multiplex detection capability [5–7].

J. Choo (✉)
Department of Bionano Engineering and Center for Integrated Human Sensing System, Hanyang University, Ansan 426-791, Korea
e-mail: jbchoo@hanyang.ac.kr

H. Chon
Department of Bionano Engineering, Hanyang University, Ansan 426-791, Korea
e-mail: hyangahchon@gmail.com

M. Baranska (ed.), *Optical Spectroscopy and Computational Methods in Biology and Medicine,* Challenges and Advances in Computational Chemistry and Physics 14,
DOI 10.1007/978-94-007-7832-0_14, © Springer Science+Business Media Dordrecht 2014

Recently, surface-enhanced Raman scattering (SERS) spectroscopy has become a promising alternative that helps to overcome various problems met in case of fluorescence detection. SERS spectroscopy using metal nanoparticles has shown promise in overcoming the low sensitivity inherent to conventional Raman spectroscopy, as it is possible to obtain an enormous Raman enhancement using hot spots on aggregated metal particle clusters. The SERS detection sensitivity is enhanced up to 10–14 orders of magnitude compared to conventional Raman spectroscopy [8–14]. Consequently, SERS detection provides a sensitivity that is comparable to fluorescence detection. Furthermore, it is possible to simultaneously detect multiple analytes using SERS because the signals are much narrower than fluorescence bands [15–18].

One of the most popular technologies for the SERS-based immunoassay uses the typical protocol of a sandwich immunocomplex platform [19, 20], with primary antibodies immobilized on a substrate with sequentially added antigens and secondary antibody-conjugated Raman probes. After the formation of a sandwich immunocomplex on a substrate, SERS signals are measured and analysed. Various kinds of biomarkers, including immunoglobulin (IgG) antigens [21–23], feline calicivirus (FCV) [24], prostate-specific antigen (PSA) [25], protein A [26], hepatitis B virus [27], and mucin protein (MUC4) [28], have been investigated using this SERS-based immunoassay technique.

To date, three different types of SERS assay platforms have been developed; gold-patterned substrates, a platform using magnetic beads, and a device platform using a microfluidic device. In this chapter, we first introduce various types of SERS nanoprobes and their bio-conjugation techniques for specific marker targeting. A novel immunoassay technique for a sensitive detection using SERS spectroscopy was investigated using three different SERS-based immunoassay platforms, which were developed for the fast and sensitive detection of biomarker proteins.

14.2 Various Types of SERS Targeting Nanoprobes for Immunoassay

Silver nanoparticles are widely used as optical enhancing agents in SERS–based immunoassays because the Raman enhancement factor is 100–1000 times greater than that of gold nanoparticles. However, silver nanoparticles have disadvantages, including short-term stability and uncontrollable size distribution. Gold nanoparticles are a more attractive choice for SERS-based immunoassays because of their favourable chemical properties and stability. In addition, aggregation is required for a large signal enhancement because the electromagnetic enhancement effects are greatly increased at the hot-spot junctions between particles. This aggregation process leads to poor homogeneity because it is nearly impossible to control the shapes and sizes of the nanoparticles. Consequently, the potential for structural control of large local fields of hot-spot junctions has attracted much interest in the fabrication, assembly, and properties of structures with nanoscale gaps.

Metal nanostructures with hollow interiors have been investigated by many research scientists. These are commonly prepared by coating the surfaces of colloidal

particles with thin layers of the desired material, followed by selective removal of the colloidal templates through calcination or wet chemical etching. These particles show strong enhancement effects from individual particles because of their ability to localize the surface electromagnetic fields through the pinholes in the hollow particle structures. Y. Xia and co-workers [29, 30] developed a simple and versatile route for the synthesis of metal nanostructures with well-defined hollow interiors using a replacement reaction between the surface of a nanoscale template and an appropriate salt precursor solution. C. E. Talley and co-workers [31] also synthesized hollow gold nanospheres (HGNs) using cobalt nanoparticles as templates and investigated their SERS application feasibility. According to their experimental results, the homogeneity of the HGNs leads to a nearly 10-fold improvement in signal consistency over standard silver substrates commonly used in SERS applications.

Encapsulated SERS probes, in which the central metal core and label molecule are encapsulated by a shell, have also been developed. For example, silica-encapsulated SERS nanoparticles were first reported by R. G. Freeman and co-workers [32] and commercialized by Oxonica. J. Choo and co-workers [33] developed silver-coated gold hybrid nanoprobes. Gold nanoparticles are superior to other nanoparticles in terms of long-term stability; however, the SERS enhancement effect of silver is 100–1000 times greater than that of gold. Thus, gold particles were coated with silver for maximum SERS enhancement.

Metal nanomaterials with different morphologies have also been used as SERS nanoprobes. Recently, rod-shaped nanoparticles have been introduced as powerful imaging probes for biomedical applications. Gold nanorods (GNRs) have two surface plasmon bands: a short-wavelength band near 520 nm and a long wavelength band in the near-infrared region. The long-wavelength band has much greater absorption intensity because of the longitudinal oscillation of the conduction-band electrons [34]. Furthermore, the gold nanorod aggregates also show stronger SERS enhancement effects than gold nanosphere aggregates because of the increase in density of junction points caused by their large surface area. The aspect ratio of gold nanorods can be controlled through the synthetic process, and it provides an opportunity to tune the surface plasmon band to obtain the maximum signal enhancement [35–38].

SERS targeting nanoprobes should always contain Raman reporter molecules to create a known source of the SERS spectrum for the molecule in question. The internal modes of the reporter molecules adsorbed to the nanoparticle surface can be used as diagnostic signals of a specific biomarker in blood serum. In particular, different reporter molecules can be applied to multiplex SERS detection. Because SERS peaks are narrower than fluorescence peaks, it is possible to create many distinct and simultaneously quantifiable SERS probes. It is also very important to functionalize the surface of the nanoparticles to selectively attach the nanoprobes to specific biomarkers in blood serum. For this purpose, antibodies or other biomolecules should be conjugated to the gold nanoparticle surface. Preparing stable biomolecule-gold nanoparticle complexes depends on several interactions: (a) electrostatic attraction between the negatively-charged nanoparticles and the abundant positively-charged sites on the biomolecule, (b) hydrophobic adsorption of biomolecules on the metal surface, and (c) covalent bonding between free thiol groups and nanoparticles. Among these conjugation methods, covalent attachment of molecules

through thiols or disulfides is the most popular choice for gold nanoparticles. Various types of biomolecules, such as antibodies, peptides, oligonucleotides, and other ligands, have been conjugated to the surfaces of metal nanoparticles using this covalent bonding.

Halas and co-workers [39] used o-pyridyldisulfide-n-hydroxysuccinimide polyethylene glycol (OPSS-PEG-NHS) to tether antibodies to the surface of gold nanoshells. The S-S terminal group of o-pyridyldisulfide was cleaved and chemically bonded to the metal surface. The other terminal group, succinimidyl propionate, was bound to an amine group on the antibody. When PEG was used, the space between the probe and antibody was enlarged, and the more favorable antibody movement greatly improved the efficiency of the antibody-antigen interaction. Choo and co-workers [40] used dihydrolipoic acid (DHLA) for the antibody conjugation. Two –SH terminal groups of DHLA were cleaved and chemically bound to the nanoparticles surface. Then, –COOH terminal groups were activated using N-hydroxysuccinimide (NHS)/1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC) coupling. Finally, antibody immobilization to the nanoparticles occurred by displacing the NHS group with lysine residues of the antibody.

J. Irudayaraj and co-workers [41] used 11-mercaptoundecanoic acid (MUDA) as the alkanethiol to react with gold nanorods to produce an activated surface for biofunctionalization. Here, antibodies were replaced using an NHS/EDC coupling reaction. M. A. El-Sayed and co-workers demonstrated gold nanorods directly conjugated to a specific peptide through a thioalkyl-triazole linker molecule. C. A. Mirkin and co-workers [42] demonstrated the detection of specific DNA sequences by tethering probe oligonucleotides to the surface of gold nanoparticles. Gold nanoparticles modified with alkylthiol-capped oligonucleotide strands were used as probes to monitor the presence of target DNA strands. Various bioconjugated gold and silver nanoparticles and their applications are well summarized and reviewed by V. Biju et al. [43].

14.3 SERS-Based Immunoassay Using Gold-Patterned Substrates

Typically, the SERS technique utilizes the sandwich immunocomplex assay. Capture antibodies are immobilized to a solid substrate, and then antigen solutions and specific Raman reporter-labelled immuno-Au nanoparticles are sequentially added. Antigen biomarkers can be quantitatively assayed by monitoring the intensity change of a characteristic SERS peak of the reporter molecule. When reporter molecules are adsorbed to roughened metal surfaces, their SERS signals are greatly increased due to electromagnetic and chemical enhancement at SERS active sites known as “hot junctions” when they are exposed to the excitation laser source. This enhancement effect shows a promise in overcoming the low sensitivity problem inherent to conventional Raman spectroscopy. Figure 14.1 displays the SERS-based immunoassay platform using gold substrates.

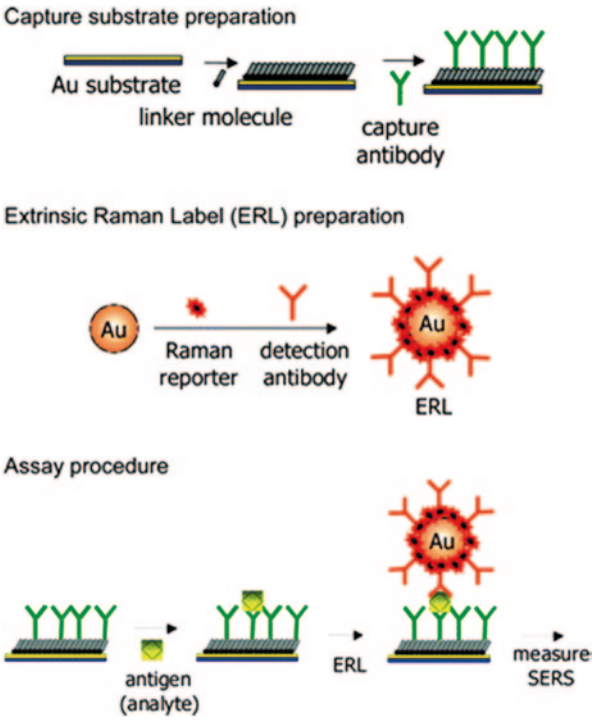


Fig. 14.1 SERS-based immunoassay platform using a planar gold substrate. Reprinted with permission from Ref. [44]

Many different types of gold substrates have been developed for SERS-based immunoassays. Porter et al. [44] fabricated a gold substrate using a coupling agent that forms a gold-bound thiolate to covalently couple with the primary amines on antibodies for immobilization. The gold substrate was prepared by evaporating 250 nm of gold onto a p-type silicon wafer. This gold-coated wafer was submerged in a solution of dithiobis(succinimidyl undecanoate) or dithobis(succinimidyl propionate) to form a thiolate adlayer on the gold substrate to covalently tether protein markers by amide linkages. Finally, this substrate was treated with a blocking agent, such as bovine serum albumin.

Choo and co-workers [19, 20] fabricated a gold-patterned microarray chip for multiplex SERS-based immunoassay. Initially, titanium metal was sputtered onto a clean glass slide, followed by gold deposition. After the photoresist was spin-coated, the gold thin film was etched to form a pattern array. Trimethoxy-propylsilane self-assembled monolayers were used as a precursor for the formation of a background hydrophobic surface. Alternatively, the opened Au wells were modified with carboxylate-terminated SAMs (11-mercaptoundecanoic acid and 6-mercapto-1-hexanol) to form a hydrophilic surface. Figure 14.2 shows the fabrication schemes of a gold-patterned microarray chip, contact angles and photographic images. A sensitive SERS-based immunoassay was performed for angiogenin, alpha-fetoprotein, and antigen IgG using this gold-patterned microarray chip.

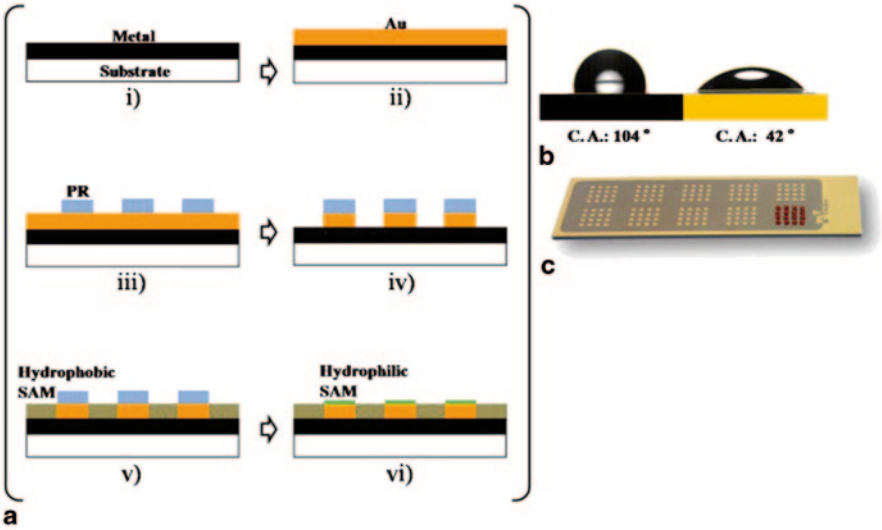


Fig. 14.2 **a** Fabrication schematics of a gold-patterned microarray chip: *i*) metal sputtering, *ii*) Au deposition, *iii*) photoresist coating, *iv*) array patterning, *v*) hydrophobic SAM treatment, *vi*) photoresist lift-off and hydrophilic SAM treatment. **b** Contact angles for the hydrophobic layer (*left*) and hydrophilic layer (*right*). **c** Illustration of a gold-patterned microarray chip. Aqueous red ink drops are selectively attached to hydrophilic wells. Reprinted with permission from Ref. [19]

Z. Tian and co-workers [45] fabricated a monolayer SERS substrate of gold nanoparticles coated with an ultrathin shell of silica (Au/SiO_2). This substrate was used instead of the Au tip traditionally applied in tip-enhanced Raman spectroscopy (TERS). Gold nanoparticles were first chemically synthesized with a core diameter of 55 nm, onto which ultrathin silica shells were coated to form Au/SiO_2 nanoparticles. Each nanoparticle acts as an Au tip in the tip-enhanced Raman system, and this technique utilizes thousands of TERS tips to the substrate surface to be probed. The main advantages of this SERS substrate are its high detection sensitivity and vast practical applications to various SERS analyses. Figure 14.3 displays the shell-isolated nanoparticle monolayer and its working principle.

14.4 SERS-Based Immunoassay Using Magnetic Beads

The SERS-based immunoassay using gold substrates mentioned above has several limitations. First, an extended incubation time was required for each binding step because the restriction on molecular diffusion near the surface slows the kinetics of protein-protein assays. Second, the repetition of washing steps to remove nonspecific binding proteins is inconvenient. Third, all of the immune-reagent components should be immobilized to a surface of a solid substrate in air. However, in many cases, the exposure of proteins to air seriously reduces their biological activity. Finally,

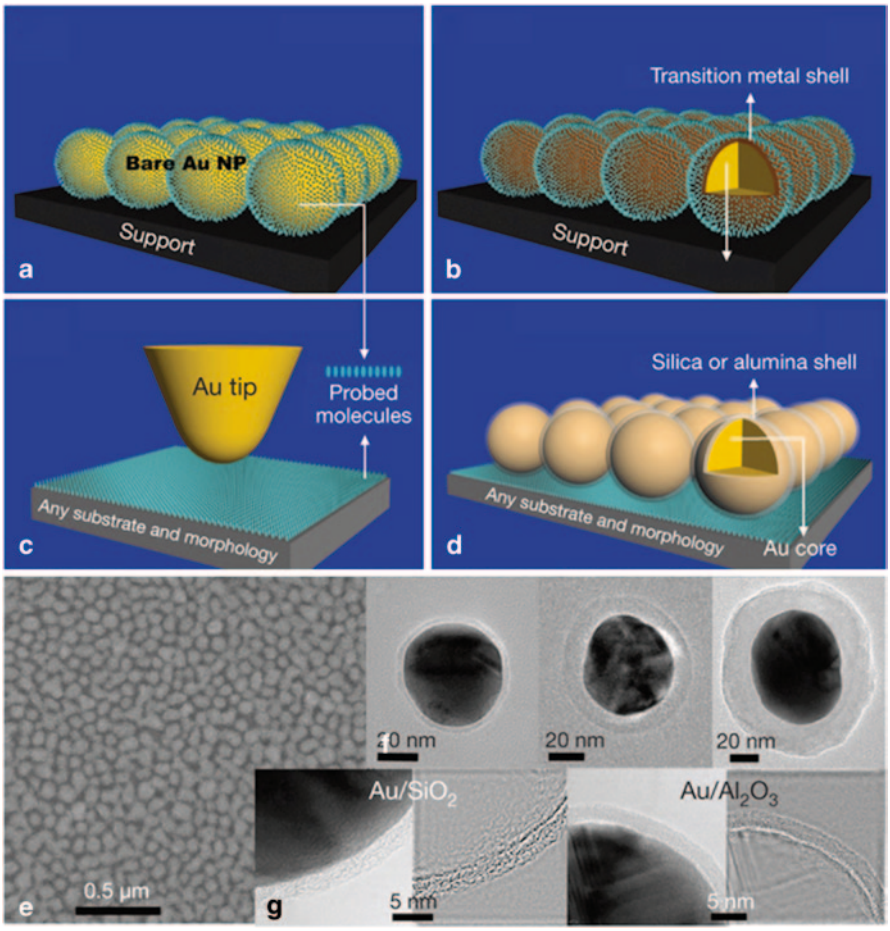


Fig. 14.3 Schematic of the contact mode. **a** Bare Au nanoparticles: contact mode. **b** Au core-transition metal shell nanoparticles adsorbed by probed molecules: contact mode. **c** Tip-enhanced Raman spectroscopy: non-contact mode. **d** SHINERS: shell-isolated mode. **e** Scanning electron microscope image of a monolayer of Au/SiO2 nanoparticles on a smooth Au surface. **f** HRTEM images of Au/SiO2 core-shell nanoparticles with different shell thicknesses. **g** HRTEM images of a Au/SiO2 nanoparticle and Au/Al2O3 nanoparticle with a continuous and completely packed shell approximately 2-nm thick. Reprinted with permission from Ref. [45]

in the immobilization-based SERS chip, it is very difficult to fabricate surfaces with highly reproducible enhancements. As a result, reproducible SERS signals could not be obtained with different laser spots.

To resolve these problems, a quick and reproducible SERS-based immunoanalysis technique using magnetic beads has been developed. This method does not use an immobilization procedure on a planar substrate; instead, it uses magnetic beads as antibody-supporting materials. In this system, the sandwich immunocomplexes are immobilized to the microtube wall using a small magnetic bar. This technique

overcomes the slow immunoreaction problems caused by diffusion-limited kinetics on a solid substrate because the reaction occurs in solution. In addition, the loading density of the capture antibody on a magnetic bead is dramatically increased because three-dimensional microspheres were used instead of two-dimensional microarrays. As a result, the assay time is greatly reduced, and the detection sensitivity is greatly increased. Furthermore, it is possible to obtain more reproducible results because the SERS signals are measured for the average nanoparticle ensemble in solution.

R. Yu and co-workers [46] demonstrated a SERS immunoassay method combining a glass-encapsulated silver nanotag as a detection probe and glass-encapsulated magnetic nanoparticles as an immobilization matrix and separation tool for an immunoassay of alpha-fetoprotein (AFP). When silver nanoparticles are used for quantitative measurements, optimal aggregation conditions are required for signal enhancement. However, in many cases, the over-aggregation of silver nanoparticles leads to particle precipitation over time, and this reduces the reproducibility of the SERS signals.

To resolve this problem, J. Choo and co-workers [47] fabricated hollow gold nanospheres (HGNs) and used them as single nanoparticle SERS agents. These particles show strong enhancement effects from individual particles because of their ability to localize the surface electromagnetic fields through pinholes in the hollow particle structures. As a result, they can be used as highly reproducible sensing probes for the quantitative immune-analysis of specific markers. The feasibility of the SERS-based detection method was investigated using HGNs and magnetic beads for a well-known lung cancer marker, carcinoembryonic antigen (CEA). Here, the LOD (limit of detection) was determined to be 1–10 pg/mL, and this value shows that SERS is approximately 100–1000 times more sensitive than an ELISA [48]. This SERS-based immunoassay technique has several advantages over previously reported SERS detection methods using the immobilization of an immunocomplex on a solid substrate with respect to good reproducibility and a fast assay time. Figure 14.4 displays schematic illustrations of HGNs, immunoassay processes, and corresponding Raman spectra.

Q. Zhang and co-workers [49] used Fe_3O_4/Au core-shell nanoparticles as SERS substrates instead of magnetic beads. The particles were prepared by combining microwave-assisted synthesis of Fe_3O_4 cores with Amal's method of coating Au shells with the aid of poly(ethyleneimine). This SERS-based immunoassay technique provides additional hot spots between gold and core-shell nanoparticles. The assay technique was applied for the detection of free prostate cancer antigens.

Z. Zhang and co-workers [50] fabricated new SERS nanoprobes with hybrid multilayered nanoshells prepared by the layer-by-layer (LBL) assembly of small silver nanoparticles on the surface of silica particles using poly(ethyleneimine). The huge electromagnetic enhancement was achieved by strong plasmon coupling between shell multilayers, and uncertain effects on SERS signals resulting from the variability of configurations are minimized by the assembly of small silver nanoparticles. Figure 14.5 shows the SERS-based sandwich immunoassay process using silica-encapsulated magnetic nanoparticles and multilayered silver shell nanoprobes.

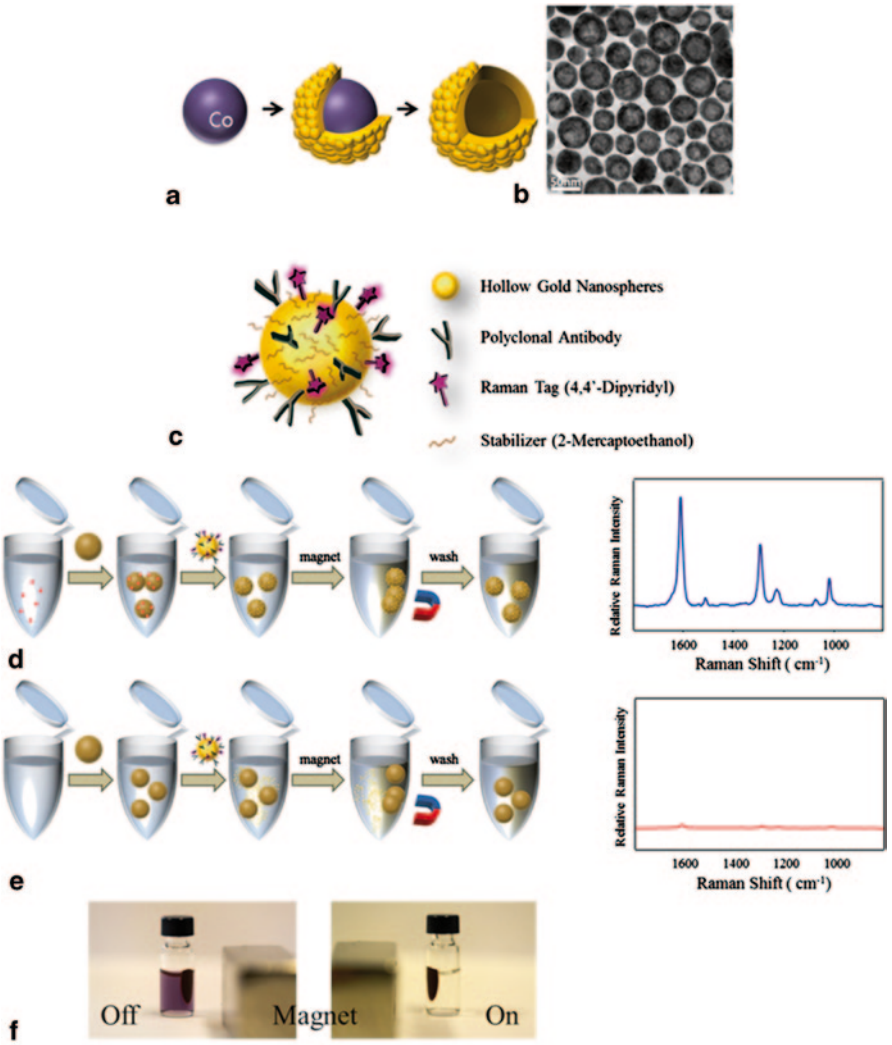


Fig. 14.4 Schematic illustration of HGNs for a SERS immunoassay: **a** experimental procedure that generates HGNs by templating against a cobalt nanoparticle. **b** TEM image of HGNs prepared using the template of cobalt nanoparticles. **c** conjugation of Raman tag (4,4′-dipyridyl) and CEA antibodies onto HGNs. Schematic illustration of immunoassay processes and corresponding Raman spectra: **d** with and **e** without CEA antigens, **f** photographs showing suspended magnetic beads attracted to the wall of a microtube by a bar magnet: with (*left*) and without CEA (*right*). Reprinted with permission from Ref. [47]

The assay was also performed for the lung cancer marker CEA, and its LOD was determined to be 0.1 pg/mL.

J. Irudayaraj and co-workers [51] performed a SERS-based immunoassay using a sandwich immunocomplex between antibody-conjugated gold nanorods and

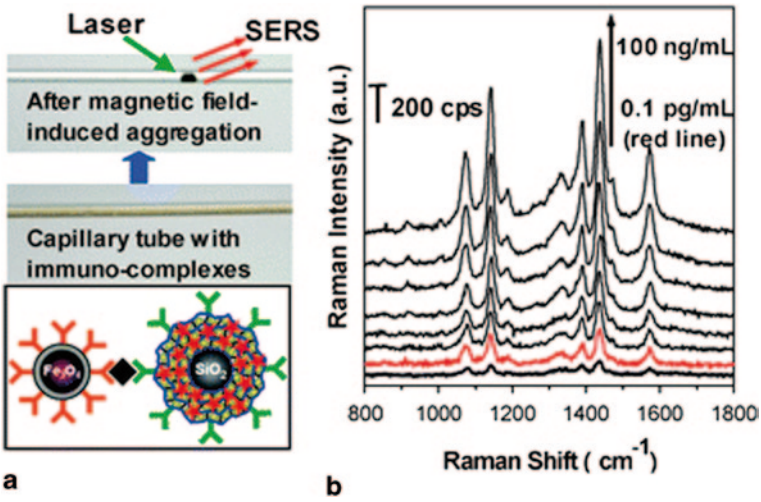


Fig. 14.5 **a** Procedure for the sandwiched immunoassay by the magnetic capture and Raman readout. **b** CEA immunoassay using the SERS nanotag-(4-ABT) in a 1:1 mixture of human serum and PBS buffer (the bottom dark line represents the blank). The concentrations are 0.1, 1.0, 10, 100 pg mL^{-1} and 1.0, 10, 100 ng mL^{-1}. Reprinted with permission from Ref. [50]

aptamer-conjugated gold nanoparticles. Taking advantage of SERS hot spots and the optical property of gold nanorods, a novel SERS aptasensor was proposed based on nanoparticle-nanorod junctions to detect human α-thrombin in human blood serum. The simulation and experiments revealed that the strong electromagnetic coupling resonance at the junction of these probes could be used to construct highly sensitive SERS aptasensors. Figure 14.6 shows a schematic representation of the SERS-based immunoassay using particle-rod junctions.

J. Choo and co-workers [52] performed a simultaneous detection of dual cancer markers in blood serum using a SERS-based sandwich immunoassay between hollow gold nanospheres and magnetic beads. For this purpose, two different Raman reporter-labelled gold nanoprobe and magnetic bead sets were prepared. The simultaneous detection of CEA and AFP cancer markers in blood serum was achieved under a single excitation wavelength. This novel immunoassay shows a simultaneous multiple-marker detection using reliable, sharp, and easily distinguishable SERS peaks. The immunoassay results for human clinical sera also show high accuracy and sensitivity. Overall, this study demonstrates a promising future for a functional nanoprobe-based SERS detection technique for a simultaneous immunoassay of multiple disease markers. Figure 14.7 displays a scheme of the SERS-based duplex immunoassay.

More recently, K. T. Carron and co-workers [53, 54] introduced a new class of SERS materials consisting of gold nanoparticles coated onto hollow, buoyant silica microspheres. These materials allow for a new molecular assay designated as a lab-on-a-bubble (LoB). These serve as a convenient platform for the detection of

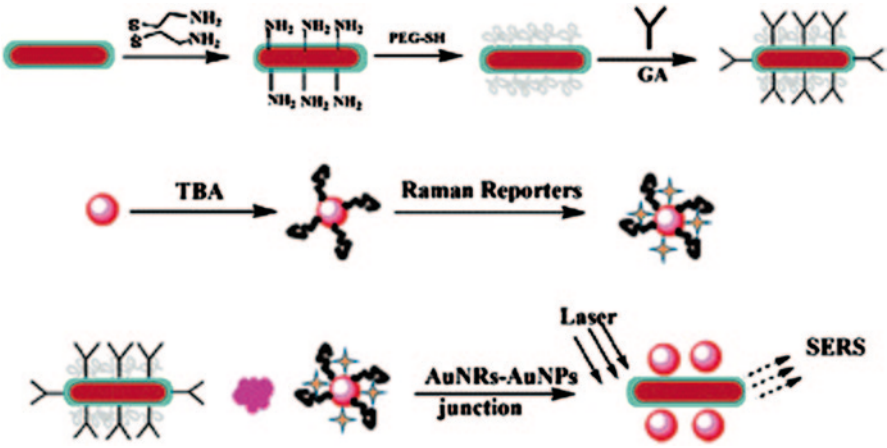


Fig. 14.6 A schematic representation of the fabrication of AuNR–AuNP junctions for protein detection. Reprinted with permission from Ref. [50]

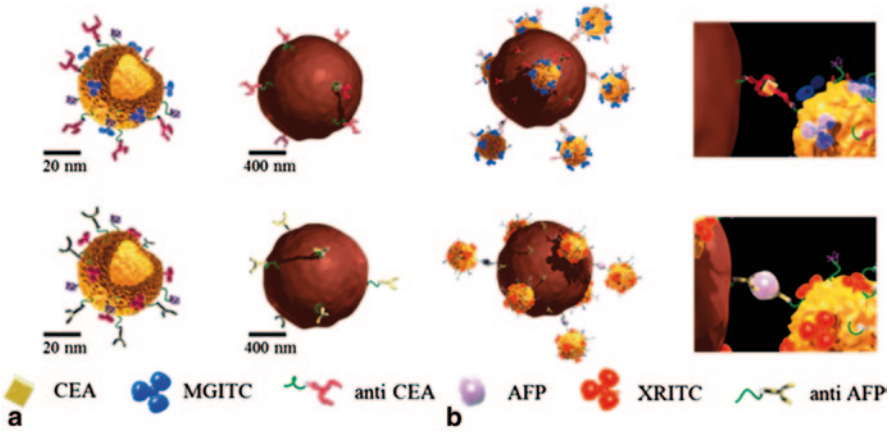


Fig. 14.7 Schematic illustration of the SERS-based duplex immunoassay using HGNs and magnetic beads. **a** Conjugation of MGITCs (Raman reporter) and CEA antibodies to HGNs (*upper left*), conjugation of XRITCs (Raman reporter) and AFP antibodies to HGNs (*upper right*), anti-CEA immobilized magnetic beads (*lower left*), and anti-AFP modified magnetic beads (*lower right*). **b** Formation of sandwich immunocomplexes between HGNs and magnetic beads for CEA and AFP. Reprinted with permission from Ref. [52]

analytes in solution and offer several advantages over traditional gold and planar SERS substrates, such as the ability to localize and concentrate analytes for detection. The LoB platform was used for a SERS-based assay of the cholera toxin. Figure 14.8 shows a schematic representation of the SERS-based sandwich assay using gold nanoparticle-coated buoyant silica microspheres.

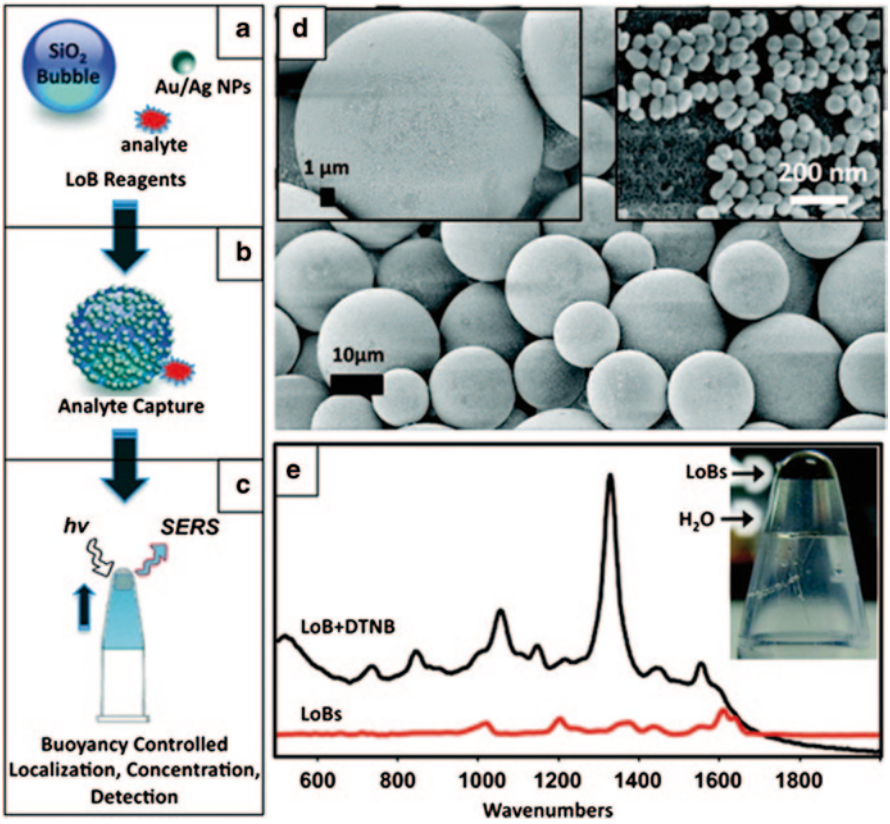


Fig. 14.8 **a–c** The basic components of a lab-on-a-bubble (LoB) assay for SERS-based detection of an analyte. **d** Representative SEM images of SERS-active AuNP-coated LoBs. **e** Representative Raman spectra of "naked" LoBs and LoBs in the presence of DTNB. The *inset* shows a picture of SERS-active buoyant LoBs in a microcentrifuge tube. Reprinted with permission from Ref. [54]

14.5 SERS-Based Immunoassay Using lab-on-a-chip

In many cases, a SERS-based immunoassay using planar gold substrates or microbeads has some issues because it is awkward to optimize the experimental conditions, such as the manual washing steps using a micropipet, changing the substrate condition for an optimal SERS measurement and the inhomogeneous distribution of sandwich immunocomplexes immobilized on the microtube wall or the planar substrate. In addition, various concentrations of a specific target marker should be manually prepared for quantitative analysis. This time-consuming process makes the substrate- or microtube-based immunoassay less attractive. Recently, SERS-based microfluidic sensors have been developed to resolve this problem. Using this platform, a highly accurate and reproducible analysis is possible if continuous

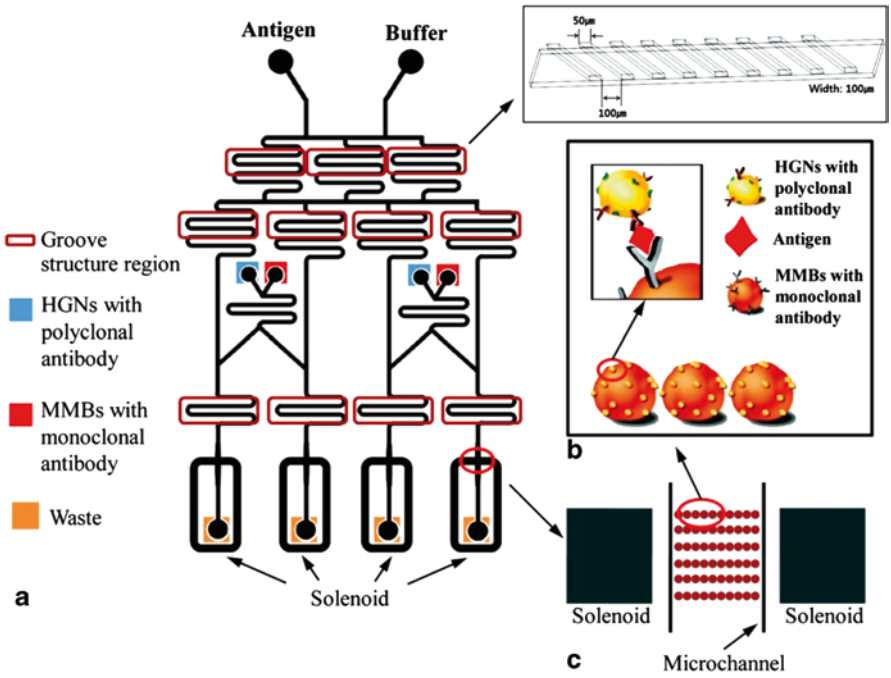


Fig. 14.9 **a** Layout of a SERS-based gradient optofluidic sensor integrated with solenoids. **b** Formation of a sandwich immunocomplex between HGNs and magnetic beads. **c** Trapping of sandwich immunocomplexes in a microfluidic channel. Reprinted with permission from Ref. [58]

flow and homogeneous mixing conditions are maintained in a microfluidic channel [55–57].

J. Choo and co-workers [58] developed a SERS-based gradient optofluidic sensor for the fast and reproducible immunoassay. In this work, a novel microfluidic sensor with functional internal structures has been designed and fabricated. This sensor is composed of three compartments consisting of the gradient channel that serially dilutes the target marker, the injection and mixing area of antibody-conjugated hollow gold nanospheres and magnetic beads, and the trapping area of sandwich immunocomplexes using multiple solenoids. This SERS-based gradient optofluidic sensor can replace the set of planar substrates or microtubes used in manual serial dilutions that have been traditionally used in SERS-based immunoassays. Figure 14.9 shows the layout of the SERS-based optofluidic sensor and its working scheme.

G. Worsley and co-workers [59] developed a multiplex lateral flow immunoassay technique using SERS-nanotags as analyte labels. Using these SERS nanotags, they performed a quantitative test for inflammation biomarkers that is transferable to point-of-care (POC) technologies. The advantages of SERS nanotags as lateral flow labels are the capability to multiplex and reduced background signal from assay components. The commercial SERS nanotags could be used in lateral flow tests

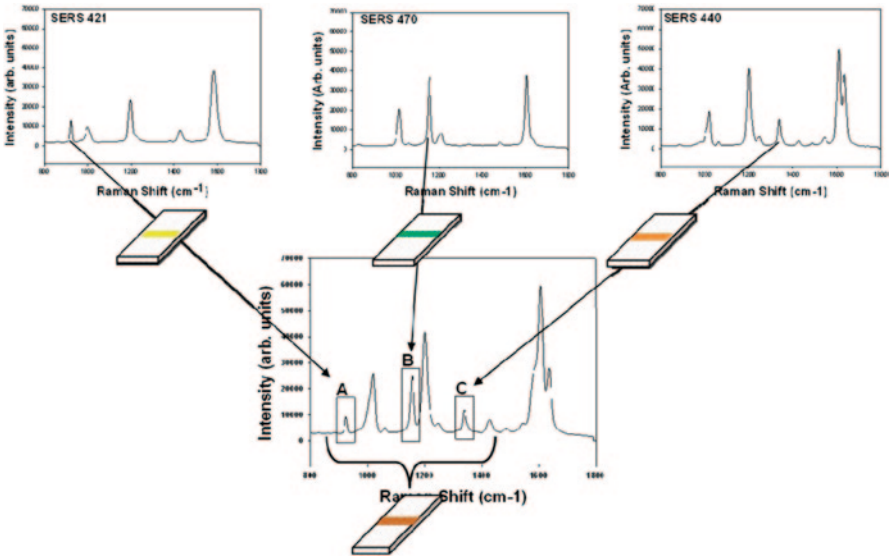


Fig. 14.10 Raman spectra of equimolar dilutions of SERS nanotags in water. The highlighted Raman peaks at (*A*) 923, (*B*) 1160, and (*C*) 1335 cm^{-1} are unique to SERS nanotags 421, 470, and 440, respectively. Scans were taken from 0 to 2000 cm^{-1}. Spectra have been cropped for clarity. Reprinted with permission from Ref. [59]

for quantitative detection, and similar sensitivity could be achieved using fluorescent labels. Figure 14.10 displays a lateral flow immunoassay using SERS nanotags.

J. Choo and co-workers [60] developed a programmable and fully automatic gold array-embedded gradient microfluidic chip that integrates the gradient microfluidic device with gold-patterned microarray wells. A serial dilution of the antigen marker can be achieved in a stepwise manner using microfluidic concentration gradient generators with N cascade-mixing stages. In this channel, the desired concentrations can be achieved by controlling volumetric mixing ratios of two merging solutions in each stage. In addition, 30-μm gold-patterned microarray wells (5 × 5) were embedded onto a glass substrate for antibody immobilization on the surface. By using this novel gold microarray-embedded gradient microfluidic channel, the tedious manual dilution process can be avoided, and highly accurate immunoanalysis could be achieved with automatic flow controls. The utility of this platform was demonstrated by the quantitative immunoassay of the alpha-fetoprotein protein marker. The total assay time, including incubation, washing steps and detection, was less than 60 min. Figure 14.11 shows the layout of a gold array-embedded gradient chip for the SERS-based immunoassay.

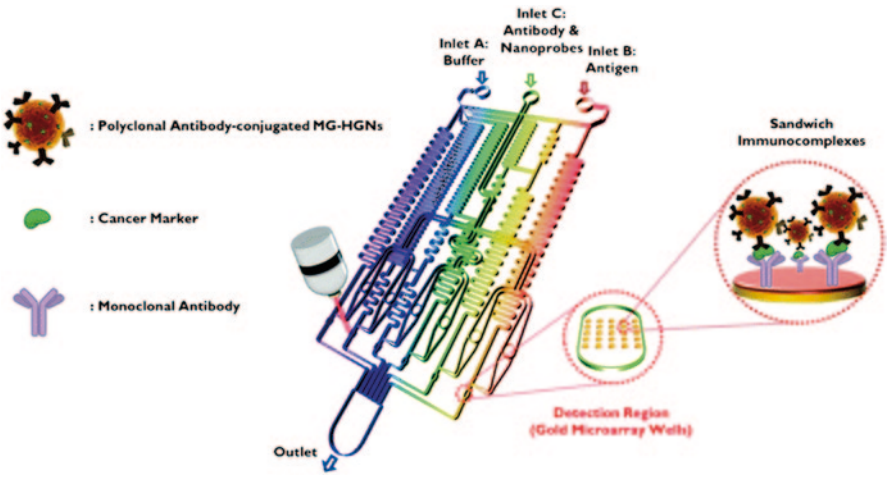


Fig. 14.11 Layout of a gold array-embedded gradient chip for the SERS-based immunoassay. The illustrations in the enlarged circles represent the formation of sandwich immunocomplexes on the surface of 5 × 5 round gold wells embedded in the gradient channel. Reprinted with permission from Ref. [60]

14.6 Summary

During the past two decades, there has been growing interest in the biomedical application of SERS spectroscopy. In particular, SERS-based immunoassays are a potential clinical tool for highly sensitive and reproducible multiple biomarker detection. In this chapter, three different types of the SERS assay platforms using (i) gold-patterned substrates, (ii) various kinds of microbead substrates, and (iii) a microfluidic device were described. In the ELISA-based immunoassay, the LOD is 1–10 ng/mL, whereas it is 1–10 pg/mL for the SERS-based assay. These LOD values demonstrate that the detection sensitivity was increased by two or three orders of magnitude compared to a conventional ELISA. Furthermore, this SERS-based immunoassay technique has additional advantages, including low reagent consumption, short incubation time, wider concentration dynamic range, and cost savings compared to the conventional ELISA technique. The SERS-based immunoassay technique can be used for highly sensitive analysis of multiple biomarkers. By using proper antibody-functionalized SERS-coding nanoprobes, the optical signals for specific immunocomplexes can be amplified and analyzed by this novel technique. As a result, this SERS-based immunoassay is expected to be a powerful clinical tool for early disease diagnosis.

Acknowledgements The National Research Foundation of Korea supported this work (grant numbers R11-2008-0061852 and K20904000004-12A0500-00410). This research was also supported by the Nano Material Technology Development Program through the National Research Foundation of Korea funded by the Ministry of Education, Science, and Technology (grant number 2012035286) and by the Agency for Defense Development through Chemical and Biological Defense Research Center.

References

1. Kanda V, Kariuki JK, Harrison DJ, McDermott MT (2004) Anal Chem 76:7257
2. Terry LA, White SF, Tigwell LJ (2005) J Agric Food Chem 53:1309
3. Rowe CA, Scruggs SB, Feldstein MJ, Golden JP, Ligler FS (1999) Anal Chem 71:431
4. Rubina AY, Kolchinsky A, Makarov AA, Zasedatelev AS (2008) Proteomic 8:817
5. Bally M, Halter M, Voros J, Grandin HM (2006) Surf Interface Anal 38:1442
6. Rosi NL, Mirkin CA (2005) Chem Rev 105:1547
7. Peterson AW, Heaton RJ, Georgiadis R (2000) J Am Chem Soc 122:7837
8. Fleischm M, Hendra PJ, Mcquilla AJ (1974) Chem Phys Lett 26:163
9. Schatz GC (1984) Acc Chem Res 17:370
10. Moskovits M (2005) J Raman Spectrosc 36:485
11. Kneipp K, Kneipp H, Kneipp J (2006) Acc Chem Res 39:443
12. Nie S, Emory SR (1997) Science 275:1102
13. Kneipp K, Kneipp H, Itzkan I, Dasari RR, Feld MS (1999) Chem Rev 99:2957
14. Kneipp K, Wang Y, Kneipp H, Perelman LT, Itzkan I, Dasari R, Feld MS (1997) Phys Rev Lett 78:1667
15. Vlckova B, Pavel I, Sladkova M, Siskova K ,Slouf M (2007) J Mol Struct 834:42
16. Kneipp K, Kneipp H (2006) Appl Spectrosc 60:322a
17. Schatz GC, Young MA, Van Duyne RP (2006) Top Appl Phys 103:19
18. Dieringer JA, McFarland AD, Shah NC, Stuart DA, Whitney AV, Yonzon CR, Young M, Zhang XY, Van Duyne RP (2006) Faraday Discuss 132:9
19. Lee M, Lee S, Lee JH, Lim HW, Seong GH, Lee EK, Chang SI, Oh CH, Choo J (2011) Biosens Bioelectron 26:2135
20. Kim I, Junejo I-u-R, Lee M, Lee S, Lee EK, Chang S-I, Choo J (2012) J Mol Struct 1023:197
21. Han XX, Cai LJ, Guo J, Wang CX, Ruan WD, Han WY, Xu WQ, Zhao B, Ozaki Y (2008) Anal Chem 80:3020
22. Li T, Guo L, Wang Z (2008) Biosens Bioelectron 23:1125
23. Ni J, Lipert RJ, Dawson GB, Porter MD (1999) Anal Chem 71:4903
24. Driskell JD, Kwarta KM, Lipert RJ, Porter MD, Neill JD, Ridpath JF (2005) Anal Chem 77:6147
25. Grubisha DS, Lipert RJ, Park HY, Driskell J, Porter MD (2003) Anal Chem 75:5936
26. Lin C-C, Yang Y-M, Chen Y-F, Yang T-S, Chang H-C (2008) Biosens Bioelectron 24:178
27. Xu SP, Ji XH, Xu WQ, Li XL, Wang LY, Bai YB, Zhao B, Ozaki Y (2004) Analyst 129:63
28. Wang G, Lipert RJ, Jain M, Kaur S, Chakraboty S, Torres MP, Batra SK, Brand RE, Porter MD (2011) Anal Chem 83:2554
29. Sun YG, Mayers B, Xia YN (2003) Adv Mater 15:641
30. Chen J, Saeki F, Wiley BJ, Cang H, Cobb MJ, Li ZY, Au L, Zhang H, Kimmey MB, Li XD, Xia YN (2005) Nano Lett 5:473
31. Talley CE, Jackson JB, Oubre C, Grady NK, Hollars CW, Lane SM, Huser TR, Nordlander P, Halas NJ (2005) Nano Lett 5:1569
32. Doering WE, Piotti ME, Natan MJ, Freeman RG (2007) Adv Mater 19:3100
33. Lee S, Kim S, Choo J, Shin SY, Lee YH, Choi HY, Ha SH, Kang KH, Oh CH (2007) Anal Chem 79:916

34. Orendorff CJ, Gearheart L, Jana NR, Murphy CJ (2006) Phys Chem Chem Phys 8:165
35. Nikoobakht B, Wang JP, El-Sayed MA (2002) Chem Phys Lett 366:17
36. Nikoobakht B, El-Sayed MA (2003) Chem Mater 15:1957
37. Nikoobakht B, El-Sayed MA (2003) J Phys Chem A 107:3372
38. Huang XH, El-Sayed IH, Qian W, El-Sayed MA (2006) J Am Chem Soc 128:2115
39. Loo C, Lowery A, Halas NJ, West J, Drezek R (2005) Nano Lett 5:709
40. Lee S, Chon H, Lee M, Choo J, Shin SY, Lee YH, Rhyu IJ, Son SW, Oh CH (2009) Biosens Bioelectron 24:2260
41. Yu CX, Nakshatri H, Irudayaraj J (2007) Nano Lett 7:2300
42. Cao YWC, Jin RC, Mirkin CA (2002) Science 297:1536
43. Biju V, Itoh T, Anas A, Sujith A, Ishikawa M (2008) Anal Bioanal Chem 391:2469
44. Porter MD, Lipert RJ, Siperko LM, Wang G, Narayanana R (2008) Chem Soc Rev 37:1001
45. Li JF, Huang YF, Ding Y, Yang ZL, Li SB, Zhou XS, Fan FR, Zhang W, Zhou ZY, Wu DY, Ren B, Wang ZL, Tian ZQ (2010) Nature 464:392
46. Gong JL, Liang Y, Huang Y, Chen JW, Jiang JH, Shen GL, Yu RQ (2007) Biosens Bioelectron 22:1501
47. Chon H, Lee S, Son SW, Oh CH, Choo J (2009) Anal Chem 81:3029
48. Hwang H, Chon H, Choo J, Park JK (2010) Anal Chem 82:7603
49. Zhou X, Xu WL, Wang Y, Kuang Q, Shi YF, Zhong LB, Zhang QQ (2010) J Phys Chem C 114:19607
50. Liu RY, Liu BH, Guan GJ, Jiang CL, Zhang ZP (2012) Chem Commun 48:9421
51. Wang YL, Lee K, Irudayaraj J (2010) Chem Commun 46:613
52. Chon H, Lee S, Yoon SY, Chang SI, Lim DW, Choo J (2011) Chem Commun 47:12515
53. Schmit VL, Martoglio R, Carron KT (2012) Anal Chem 84:4233
54. Schmit VL, Martoglio R, Scott B, Strickland AD, Carron KT (2012) J Am Chem Soc 134:59
55. Park T, Lee S, Seong GH, Choo J, Lee EK, Kim YS, Ji WH, Hwang SY, Gweon DG, Lee S (2005) Lab Chip 5:437
56. Lee S, Choi J, Chen L, Park B, Kyong JB, Seong GH, Choo J, Lee Y, Shin KH, Lee EK, Joo SW, Lee KH (2007) Anal Chim Acta 590:139
57. Quang LX, Lim C, Seong GH, Choo J, Do KJ, Yoo SK (2008) Lab Chip 8:2214
58. Chon H, Lim C, Ha SM, Ahn Y, Lee EK, Chang SI, Seong GH, Choo J (2010) Anal Chem 82:5290
59. Noble J, Attree S, Horgan A, Knight A, Kumarswami N, Porter R, Worsley G (2012) Anal Chem 84:8246
60. Lee M, Lee K, Kim KH, Oh KW, Choo J (2012) Lab Chip 12:3720

Chapter 15
FTIR Imaging of Tissues: Techniques and Methods of Analysis

Kamilla Malek, Bayden R. Wood and Keith R. Bambery

Abstract In this chapter, we describe biomedical applications of infrared microscopic imaging applied to human tissue sections. The central focus is human diseases including cervical cancer, neurodegenerative pathologies, and dysfunctions of cardiac and liver tissues. In addition, we briefly describe the fundamentals of FTIR imaging instrumentation along with spectral pre-processing and hyperspectral image reconstruction. The chapter concludes with a summary of what is required to take FTIR imaging technology into the clinical environment.

Keywords FTIR imaging • Data processing • Cervical cancer • Heart disease • Neurodegenerative pathologies • Liver diseases

15.1 Introduction

Fourier Transform infrared (FTIR) spectroscopic imaging is set to become a true independent modality for the diagnosis of disease. The spectroscopic advantage lies in the fact that the chemical change must precede or accompany any morphological change that is symptomatic of disease [1]. The ability to spectroscopically analyse and spatially locate macromolecules within single cells and tissues offers a platform to investigate, diagnose and monitor the treatment of neoplasia and other diseases.

In terms of diseases cancer grading by light microscopy is notoriously subjective and highly dependent on the training and experience of the pathologist examining the tissues, as well as the sample quality and quantity. Many neoplasias require an-

K. Malek (✉)
Faculty of Chemistry and Jagiellonian Centre for Experimental Therapeutics,
Jagiellonian University, Ingardena 3,
30-060, Krakow, Poland
e-mail: malek@chemia.uj.edu.pl

B. R. Wood
Centre for Biospectroscopy and School of Chemistry, Monash University,
3800, Victoria, Australia

K. R. Bambery
Australian Synchrotron, 800 Blackburn Road,
Clayton, VIC, 3168, Australia

M. Baranska (ed.), *Optical Spectroscopy and Computational Methods in Biology and Medicine,* Challenges and Advances in Computational Chemistry and Physics 14,
DOI 10.1007/978-94-007-7832-0_15, © Springer Science+Business Media Dordrecht 2014

cillary diagnostics (e.g., IHC, flow cytometry, PCR) to further characterize their cell origin, clonality and/ or malignant risk, and these tests are often time-consuming, laborious, and costly. Cancer research has focused intently on identifying mutations and other biomarkers to help detect neoplasias earlier, or to develop treatments [2–5]. However, disease phenotypes are typically extremely complex, and the examination of a single gene mutation for instance, does not adequately characterize the disease [6]. Metabolomics, in comparison to proteomics or genomics, can quantify many metabolites in complex biological systems [7]. Metabolism is more closely associated with an organism's phenotype [8] and will almost certainly be affected by disease, so metabolic biomarkers are likely to be accurate indicators of disease. Fourier transform infrared spectroscopy and Raman spectroscopy are examples of newer applications used in metabolomics. FTIR has proven to be a sensitive analytical tool to examine the structure and conformation of lipids, proteins and nucleic acids and also of complex biological materials such as body fluids and cell cultures in pursuit of biomarkers. We define an "infrared biomarker" as a region within a complex mid-infrared spectrum that is diagnostically useful. The ability of FTIR spectroscopy to detect changes in the chemical composition and morphology of cells enables its use in distinguishing differences among cell populations, such as normal cells *versus* neoplastic cervical tissue. Given the ease with which infrared spectroscopy can be performed on small tissue or cell samples with high precision and minimal preparation, an infrared parameter that relates to cellular biosynthetic activity may be valuable as a prognostic indicator, particularly in cases where the current "gold standards" are wanting.

The adaptation of multi-channel infrared array detectors from military hardware to FTIR microscopes in the early 1990s resulted in new methodologies to investigate the macromolecular architecture of cells in tissue sections [9]. The new generation of focal plane array (FPA) and linear array detectors are capable of recording thousands of spectra in rapid time. Each FTIR image is essentially a digital hyperspectral data cube containing absorbance, wavenumber and ***x, y*** spatial coordinates. Multivariate imaging techniques including UHCA and K-means clustering, Principal Components Analysis (PCA), Linear Discriminant Analysis (LDA), Fuzzy C-means clustering, and neural networks have proven to be invaluable in the identification and correlation of spectral groups or "clusters", which can be directly compared to stained cells for traditional light microscopic evaluation by pathologists.

A number of exciting new developments have also emerged in the field of FTIR spectral processing and image reconstruction. These include algorithms to correct for light scattering and determine end member spectra in complex data sets. Over the past few years, we have seen developments in the understanding of sample preparation and the choice of sample substrate has also become a hot topic especially in regard to trans-reflection based substrates, because an electric-field standing wave is produced in a transflection measurement, which causes a non-linear change in absorbance with sample thickness [10]. New insights into the interpretation of spectral bands and the mystery as to why the DNA bands appear so weak in the spectra of fixed cells and tissues have now been resolved [11, 12].

The stage is now set for the transplantation of the FTIR imaging spectrometer from the laboratory to the clinic. This chapter provides an overview and update on the progress in the field of FTIR imaging applied to cells and tissues. The chapter starts with a summary on band interpretation in cell and tissue spectra before discussing instrumentation, image data processing, and finally moving onto some case studies of different diseased tissues including cervical, brain, liver, and heart.

15.2 Spectral Features of Macromolecules in FTIR Spectra of Cells and Tissues

Fourier transform infrared spectrum (FTIR) of any biological system such as a cell or a tissue is a fingerprint of the phenotypic expression in response to environmental stimuli. It expresses the overall chemical information that highlights first of all the presence of the major biocomponents such as proteins, lipids, carbohydrates, and nucleic acids. Here, it should be stressed that FTIR spectroscopy is not a technique that enables the detection of an individual chemical species that appears as a metabolic product or a stress marker but it is rather a method to monitor the entire biochemical status of the biological processes. The change of spectral biomarkers specific for the given group of biomolecules can indicate the presence of a particular disease, environmental stress or genetic modifications. Such a "fingerprint" of biochemical status can provide the first hint about the observed biological or pathological processes; however, the comprehensive understanding of FTIR spectra requires the use of advanced analytical methods such as a multivariate analysis. This is discussed in detail in Sect. 4.

Figure 15.1 shows FTIR spectra of individual macromolecules present in a sample of a biological origin and a spectrum typical for a cell or a tissue. The comparison of band position and shape in all the spectra clearly shows that bands originating from the most abundant biocomponents are observed in the spectrum of a cell or a tissue. The second derivative IR spectrum presented in Fig. 15.2 emphasizes small differences in spectral band shapes and illustrates the dominant absorption features for proteins, lipids, carbohydrates, and nucleic acids. We discuss these in detail below, while Table 15.1 summarizes the assignment of IR bands observed in FTIR spectra of cells and tissues to particular vibrations and macromolecules.

15.2.1 Proteins

Proteins are the dominating constituents in all cells and tissues, and thus their vibrational fingerprint is easily observed in infrared microspectroscopy. The protein IR spectrum has two primary features, the amide I and amide II bands, which are

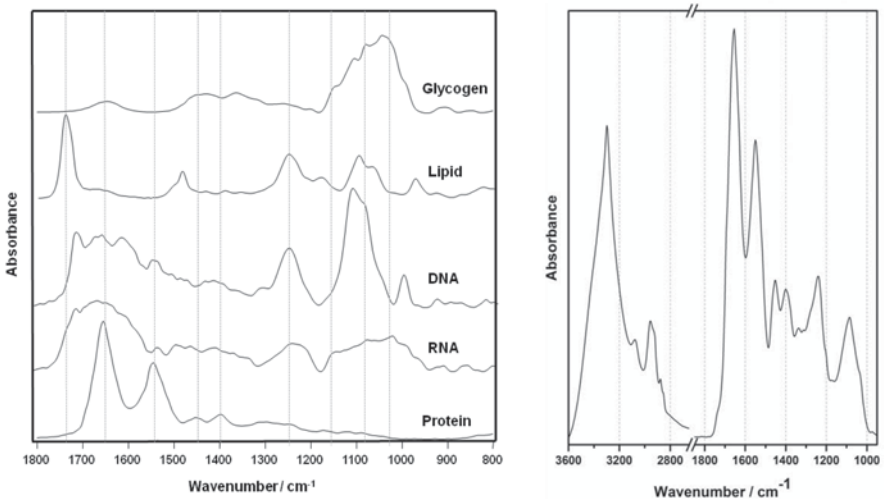


Fig. 15.1 FTIR spectra of major biological components revealing the characteristic spectral regions for proteins, nucleic acids, lipids, and carbohydrates (glycogen) (*left*); a typical FTIR spectrum of a sample of a biological origin (*right*)

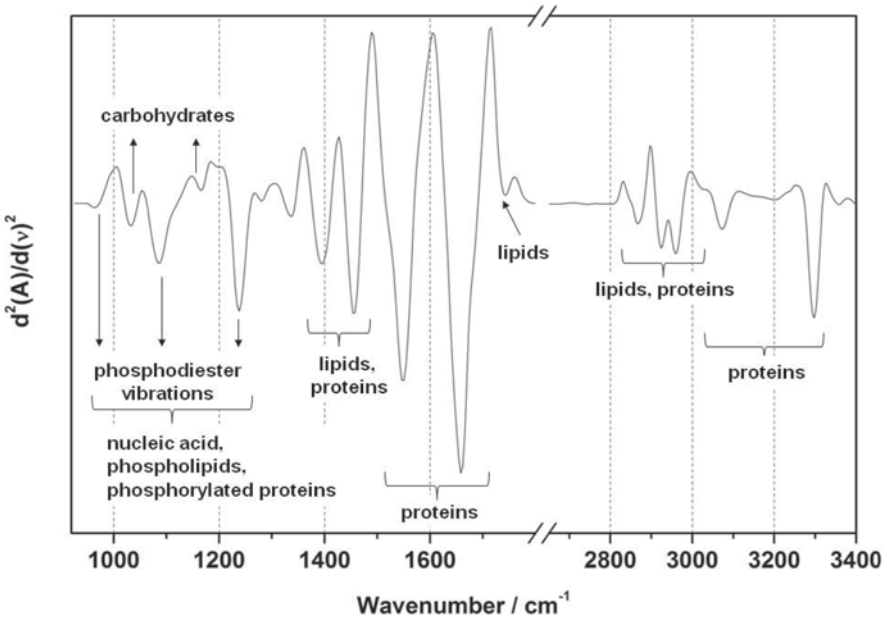


Fig. 15.2 Second derivative IR spectrum of a sample of a biological origin with labelled spectral regions specific for macromolecules

Table 15.1 Assignment of FTIR bands of cells and tissues

Vibrational mode	Wavenumber (cm^{-1})	Assignment
NH stretch (amide A)	ca. 3290	Proteins
NH stretch (amide B)	3060	Proteins
=CH olefinic stretch	3010–3030	Unsaturated lipids
CH_3 asymmetric stretch	2954–2962	Lipids and protein
CH_2 asymmetric stretch	2915–2924	Saturated lipids and side chain of proteins
CH_3 symmetric stretch	2870	Lipids and protein
CH_2 symmetric stretch	2847–2855	Saturated lipids and side chain of proteins
SH stretch	2550	Cysteine thiol group
C=O stretch (ester)	1740–1730	Lipids, phospholipids
C=O stretch	1705–1690	Bases of nucleic acids
C=O stretch+NH bend (amide I)	1695–1670	Proteins—anti-parallel β-sheet and β-turn structures
	1670–1660	Proteins—3_{10}-helix structure
	1660–1648	Proteins—α-helix structure
	1637–1623	Proteins—β-sheet structure
ring CC stretch of Tyr and Phe residues, and nucleotides	1620	Proteins, nucleic acids
NH bend+C-N stretch (amide II)	1580–1480	Proteins
ring CC stretch of Tyr residues	1520	Tyrosine proteins
CH_2 scissoring mode	1473–1462	Lipids proteins,
CH_2, CH3 deformation modes	1455	Proteins
COO^- symmetric stretch	1400	Fatty acids, amino acids
CH_3 deformation of aliphatic side groups of amino acid residues	1379	Proteins
CH_2 wagging	1343	Phospholipids, fatty acids, triglycerides, amino acid side chains
Amide III	1340–1240	Proteins
PO_2^- asymmetric stretch	1238	DNA
CO-O-C asymmetric stratech	1171	Cholesteryl esters
C-OH stretch	1171	Seronine, thyrosine, threonine
C-O stretch	1151	Glycogen, mucin
C-O stretch of ribose ring	1120	RNA
PO_2^- symmetric stretch	1083–1086	Nucleic acids, phospholipids
C-C stretch	1078	Glycogen
-CO-O-C stretches	1063	Cholesterol esters, phospholipids
COH deformation	1050–1056	Mucin, carbohydrates
	1030	Nucleic acids
	1022–1028	Glycogen, carbohydrates
Dianionic phosphate monoester	970	Phosphorylated proteins
C-N-C stretch of ribose-phosphate skeletal vibrations	965	Nucleic acids
	933	Z type DNA

widely used in studies on protein secondary structure since both of these are sensitive to structural configuration. These vibrations assigned predominantly to the C=O stretching and NH bending modes, respectively, are observed in the 1670–1620 and 1560–1530 cm^{-1} regions. Other bands characteristic for proteins only give maxima in the high-wavenumber region above 3020 cm^{-1} and they are attributed to the NH stretching vibrations (amide A and B). However, amide A and B are not usually treated as spectral biomarkers. The amide I mode is most extensively used to quantify α-helical, β-sheet, turn, and unordered structures of proteins. Usually, IR spectra of cells and tissues exhibit amide bands at *ca*. 1665 and 1540 cm^{-1} that signify protein α-helical chains. The appearance of shoulders at *ca*. 1695–1670 and 1640–1620 cm^{-1} indicates, in turn, changes in protein secondary conformation towards a β-sheet structure. For instance, such shifting was observed in IR spectra of red blood cells from β-thalassemia patients [13] and during spontaneous differentiation of murine embryonic stem cells in their early development [14], while β-amyloid peptides were detected in Alzheimer's diseased tissue [15]. On the other hand, the appearance of a band at 1624–1630 cm^{-1} was also correlated with the protein nitration process in the atherosclerotic lesion formation [16]. Shifting of the amide I band should be carefully interpreted, because dispersion artefacts can appear in this spectral range. In such cases, the wavenumber value of this band can be randomly shifted *ca*. 15 cm^{-1} and its absorbance is often reduced [17].

A number of less pronounced bands from proteins appear in the region of 1460–1240 cm^{-1}, however, vibrations of other biomolecules can contribute to these bands (see Table 15.1 for details). In the region of the scissoring CH_3 mode, a band at 1397 cm^{-1} is assigned to proteins with some contribution from lipids, whereas a band at 1400 cm^{-1} is attributed to the symmetric stretching mode of the free carboxylate group of amino acids as well as fatty acids. In some cases, IR bands highlight the presence of proteins containing Tyr and Cys residues. The breathing vibration of the tyrosine ring appears at *ca*. 1520 cm^{-1}, while the stretching mode of the cysteine thiol group is assigned to a weak and broad band with a maximum around 2550 cm^{-1}. Both of these are rarely observed in raw IR spectra but are observed in the second derivative spectra [13, 18]. In turn, the stretching C-OH mode of serine, threonine, and tyrosine were found at 1171 cm^{-1} in FTIR spectra of malignant oesophageal and neoplastic cervical tissues due to the absence of the glycogen band normally observed at 1155 cm^{-1} in normal tissue [19, 20].

FTIR spectra of some tissue sections exhibit markers characteristic of connective tissue. This tissue mainly consists of proteins (e.g. collagen) with a small component of nucleic acids. A typical feature for the connective tissue is a triad of absorbances in the region of the amide III vibrations, i.e. 1206, 1238, and 1280 cm^{-1} originating mainly from collagen. It is easily recognized by the positions of the peaks and the relative intensities, which are accompanied by a broad amide II band at 1550 cm^{-1} [21]. In turn, cells in malignant tissue often exhibit an increase in the intensity of the 970–965 cm^{-1} band, which is generally correlated to an increase in the relative content of nucleic acids. However, phosphorylation of newly formed proteins, for example, in the process of carcinogenesis [19] can also contribute to this band. A similar band was also observed when applying synchrotron IR mea-

surements to PC12 cells during neuronal differentiation, which was correlated with phosphorylation using fluorescence microscopy [22]. It has been reported that an increase in the intensity of the bands at 1080 and between 990 and 970 cm^{-1} reflects sustained elevated phosphorylation of proteins, while the asymmetric phosphate mode [$\nu_{as}(PO_2^-)$], which is observed at *ca*. 1238 cm^{-1}, appears to be less sensitive and less specific to this process [22].

An analysis of FTIR spectra of cells and tissues often includes monitoring of changes in selected bands ratios. Benedetti and co-workers showed that the ratio of amide I/amide II bands can serve as a spectral marker for changes in DNA concentration [23]. Since this ratio is equal to unity in human erythrocytes that are anucleated, its increase can indicate DNA synthesis because of the contribution of the DNA C=O vibration to the amide I absorbance (spectral features of nucleic acids are discussed in Subsections 2.4 and 2.5). In addition, the ratio of the integrated absorbances of bands at 1080 ($\nu_s(PO_2^-)$) and 1540 cm^{-1} (amide II) was proposed as a clinical parameter for the evaluation of the degree of malignancy in patients affected by B-chronic lymphatic leukaemia [23]. A similar observation was reported for rabbit bone marrow fibroblasts transformed by murine sarcoma virus [24]. In the case of IR imaging of cells, the amide I/$\nu_s(PO_2^-)$ ratio gives an insight into the cytoplasm to nucleus ratio (C/N), since a high protein content is expected to be in the cytoplasm region of the cell, while nucleic acids are condensed in the cellular nucleus [25].

15.2.2 Lipids

High lipid content mainly arises from plasma membranes and microsomal pellets such as endosomes, secretory vesicles, while only a low lipid concentration is found in cytoplasm. The most discernible spectral features of lipids or phospholipids originate from the stretching modes of the chain CH_2 and=CH groups (2800–3050 cm^{-1}), the ester C=O bond (*ca*. 1736 cm^{-1}) as well as from the scissoring vibrations of the long-chain CH_2 groups (*ca*. 1468 cm^{-1}), *c.f.* Table 15.1. Biochemical processes, in which lipid metabolism is strongly involved, are mostly recognized by the examination of the high-frequency region. There are usually only four bands observed from lipids, which are attributed to the asymmetric and symmetric stretching vibrations of the methyl and methylene groups that are found at ~2960 [$\nu_{as}(CH_3)$], 2923 [$\nu_{as}(CH_2)$], 2873 [$\nu_s(CH_3)$], and 2850 cm^{-1} [$\nu_s(CH_2)$]. In addition, a weak band at *ca*. 3010–3030 cm^{-1} is assigned to the CH stretching vibration of unsaturated fatty acids [ν(=CH)]. The bands of the methylene group are mainly due to the saturated chains in lipids, whereas the $\nu(CH_3)$ bands are due to vibrations of the methyl groups in proteins, lipids, and nucleic acids. The most pronounced spectral changes indicating alternation of lipid structure and content are found in both the methylene vibrations [26]. An increase in the ratio of the methylene bands relative to the $\nu_{as}(CH_3)$ band along with the lack of changes in intensity of the ν(=CH) band shows the presence of predominately long chain saturated fatty acids, whereas an increase of the $\nu_s(CH_3)/\nu_s(CH_2)$ ratio is indicative of branched chain fatty acids. If

there is only a low intensity $\nu_s(CH_2)$ band and no C=O band at 1740 cm^{-1} then these bands are associated with aliphatic functional groups of proteins. For instance, the high-frequency region was shown as a potential marker for cytoplasmic and nuclear changes during oocyte maturation at the germinal vesicle and metaphase II stages [26]. In addition, the position of the symmetric stretching mode of the CH_2 groups can be an indicator of order and disorder of the lipid-membrane, which is strongly dependent on the composition, hydration, content of membrane proteins and other factors [27]. In such cases, a useful parameter can be the shift in the wavenumber of the CH_2 stretching vibration [26] that indicates changes in acyl chain flexibility. The shifting towards lower frequencies can also indicate an increase in the number of *trans* isomers of lipids [28] or in general a different lipid environment. The former can be related to an increase in lipid order and chain rigidity.

Furthermore, the variability of intensities of the asymmetric CH_2 and the ester C=O stretches in relation to amide I/II bands is also an indicator of the overall protein/lipid ratio. Changes of this ratio due to lipid-protein metabolism also result in spectral changes in the region below 1500 cm^{-1}. A band at 1397 cm^{-1} that is assigned to the bending mode of the side chains of amino acid residues exhibits a higher intensity for high-content of proteins in comparison to the scissoring mode of the CH_2 groups observed at around 1460 cm^{-1} (mostly lipids).

An analysis of the ν(C=O) absorbances can also discriminate structure and composition of mitochondrial and plasma membranes in cells [27]. The C=O stretching mode at *ca.* 1740 cm^{-1} is attributed to triglycerides as well as to cholesterol esters [29], while a band at 1713 cm^{-1} was proposed as evidence for the presence of free fatty acids [16]. However, this band is usually overlapped by the amide I band. Intensity alteration of the 1470–1460 cm^{-1} band suggests changes in content of ester-containing macromolecules, which can be supported by similar changes observed in the intensity of bands at 1343 (the CH_2 wagging mode) and 1170 cm^{-1} (the CO-O-C asymmetric stretches) originating from triglycerides and cholesterol esters, respectively. In turn, the C-O vibrations of cholesterol found in atherosclerotic aortas, contribute to a band at 1055 cm^{-1} [29]. The presence of phospholipids *versus* lipids can be recognized additionally by the appearance of the phosphate group bands, especially at *ca*. 1088 cm^{-1}. However, this band is often superimposed by sub-bands originating from phosphodiester groups of various biomolecules [27].

15.2.3 Carbohydrates

In general, the spectral range of 1150–1000 cm^{-1} is typical for carbohydrates. A maximum at *ca.* 1085 and 1050 cm^{-1} with a shoulder at around 1030 cm^{-1} is assigned to mono- and polysaccharides as well as proteoglycans. The presence of a particular carbohydrate is usually linked to known information about biochemical events. Its spectral profile depends on hydration but Chiriboga and co-workers reported that for example glycogen in the cell is in a semi-hydrated form and thus its characteristic bands are observed at *ca*. 1150, 1080 and 1028 cm^{-1} [30]. The

Table 15.2 Characteristic IR bands of RNA [23, 34]

Vibrational mode	Wavenumber (cm^{-1})
C=O stretch	1705–1690 (m)
C=C stretch of imidazole ring	1610 (s)
C=N stretch of imidazole ring	1578 (m)
PO_2^- asymmetric stretch	1244 (m)
CH_2 deformation mode	1218 (w)
C-O stretch of ribose	1160 (w, sh), 1120 (s, sh)
PO_2^- symmetric stretch	1084 (s, sh)
C-O stretch of ribose	1060 (s), 1050 (s, sh); 1038 (s, sh); 1015 (m, sh)
Ring stretch and deformation of uracyl	996 (m)
Ribose-phosphate skeletal vibrations	970, 916, 880, 870, 815 (w)

s strong, *m* medium, *w* weak, *sh* shoulder

presence of the glycogen spectral features was used for discrimination of normal epithelial cells of the cervix and the oesophagus from those with cancerous features [19, 31]. The glycogen bands at 1080 and 1028 cm^{-1} may often obscure the nucleic acids bands and can confound disease diagnosis. In turn, IR maxima at 1150 and 1050 cm^{-1} found in spectra of the oral mucosa, cervical cells, and cytoplasm of glandular cells were assigned to mucin consisting of glycosylated proteins [32, 33]. In fact, mucus containing mucin covers the surface of cells, while glandular cells secrete mucin. The 1050 cm^{-1} band of mucin can mask the 1038 cm^{-1} band attributed to DNA and RNA.

15.2.4 Ribonucleic Acid (RNA)

FTIR spectra of nucleic acids show in general a number of characteristic bands. As expected, they originate from modes of the acid constitutes, i.e. the C=O and ring CC stretching modes of the purine and pyrimidine bases (in the region of 1720–1600 cm^{-1}), the asymmetric and symmetric stretches of the phosphate group as well as stretches of the C-O bonds in the phosphodiester and sugar moieties (in the region of 1300–900 cm^{-1}). For a dry sample of a cell or a tissue, the most characteristic bands appear in the entire fingerprint region, see Table 15.1. However, IR bands of nucleic acids in the region of 1720–1600 cm^{-1} are overlapped by protein bands that always appear in samples of animal/human origin. Therefore, the presence of nucleic acids is predominantly revealed by IR bands in the 1300–900 cm^{-1} region (c.f. Fig. 15.1), and are dramatically affected by environmental factors and are only observed in the absence of glycogen [27]. The spectral profiles of both DNA and RNA are significantly distinct, making FTIR analysis a useful tool for monitoring structural and content modifications of these principal constituents of cells and tissues. The characteristic IR bands of pure RNA with their assignment are collected in Table 15.2.

Benedetti and co-workers compared FTIR spectra of dried films from calf thymus DNA and RNA from baker's yeast [23]. They showed that the asymmetric [$\nu_{as}(PO_2^-)$] and symmetric [$\nu_s(PO_2^-)$] stretching modes of the phosphate group are observed at *ca.* 1240 and 1080 cm^{-1}, respectively, in the IR spectra of both nucleic acids. However, the $\nu_{as}(PO_2^-)$ mode of pure RNA was reported to be at a higher wavenumber value than that of DNA. The appearance of the two bands in IR spectra of cells and tissues can be also attributed to phosphorylated proteins, phospholipids, and other phosphate-containing macromolecules, as discussed above. In turn, IR bands at 1160, 1120, 1036, 996, and 916 cm^{-1} with their bandwidth are specific for RNA. Among them, the pronounced 1120 and 996 cm^{-1} bands assigned to vibrations of the ribose and uracyl moieties, respectively, provide unique markers to detect RNA. Although, vibrational modes of the carbohydrate CH_2OH and C-O groups also appear in this spectral region, they are less intense than nucleic acids bands [34].Increasing concentration of RNA considerably modifies the most characteristic DNA band centred at *ca*. 1240 cm^{-1}. As mentioned in paragraph 2.1, increasing DNA/RNA content also alters the protein profile in the region of the amide I and II bands. Broadening of the amide I band is greater for RNA than for DNA, and additionally shoulders at 1710 and 1705 cm^{-1} are observed in IR spectra of protein-nucleic acid mixtures with a high concentration of DNA and RNA, respectively [23].

In turn, Andrus and Strickland proposed cancer grading based on increasing absorbance ratio of bands at 1121 and 1020 cm^{-1} assigned to RNA and DNA vibrations, respectively [35]. A value of this index indicates clinically aggressive and clinically indolent lymphomas and it rises with lymphoma grade showing increasing RNA content. A similar observation was found in IR spectra of cervical cancer tissues, suggesting a higher transcription rate during carcinogenesis [36].

Another example of RNA detection in monitoring biochemical events in cells by the use of FTIR microspectroscopy was reported by Ami and co-workers [14]. In this study, the low-frequency bands of RNA, i.e. at 994 and 914 cm^{-1}, along with a DNA band at 966 cm^{-1} were identified as marker bands for an analysis of conformational and content alternations of nucleic acids in the differentiation process of embryonic stem cells. Changes in intensity of these bands revealed a low content of DNA and a high content of RNA at beginning of differentiation. The latter decreased between 4 and 7 days with shifting of the protein amide I band. This was interpreted as the beginning of mRNA transcription, in which dormant mRNA could be polyadenylated and activated to produce a new phenotype of proteins. In turn, the appearance of DNA bands at this stage of differentiation was associated with gene transcription, in which a DNA/RNA hybrid is formed [14].

15.2.5 Deoxyribonucleic Acid (DNA)

The main DNA vibrations appear at ~1225 cm^{-1} and 1080 cm^{-1} and are assigned to the phosphodiester asymmetric and symmetric stretching vibrations, respectively. The position of the asymmetric phosphodiester stretching vibration in DNA is

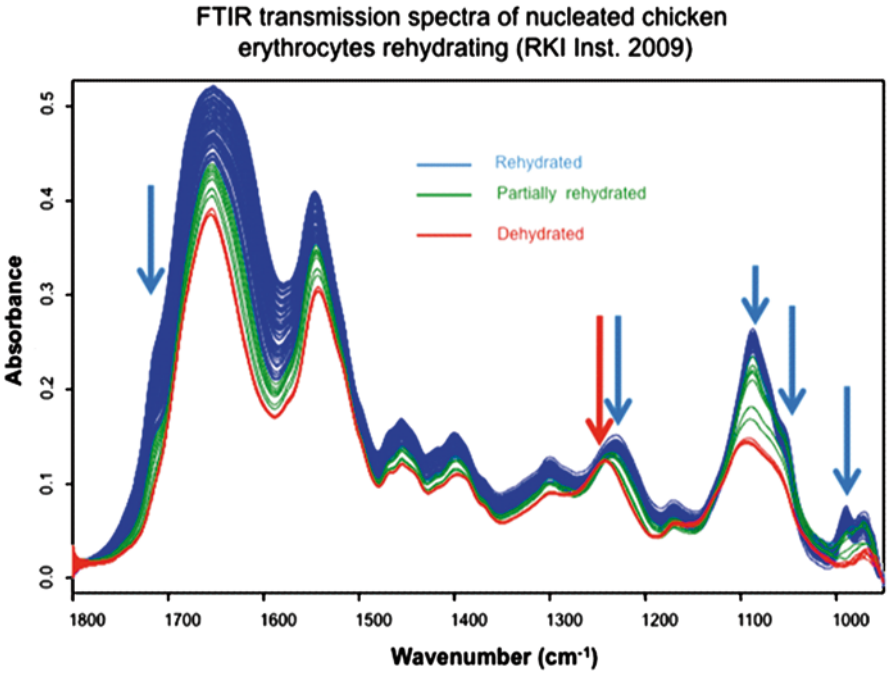


Fig. 15.3 FTIR transmission spectra recorded of chicken erythrocytes as a function of hydration. The *arrows* highlight the DNA bands that radically change during hydration

dependent on the conformation of the DNA. In low ionic strength solutions, native DNAs and most synthetic double-strand poly- and oligonucleotides adopt a B-form conformation. Under the same conditions double-stranded RNAs adopt an A-type conformation [37]. In the A-DNA conformation, which is the case for all dehydrated cells and tissues, the band appears at ~1244 cm^{-1}, while in the B-DNA conformation, which is the case for live and hydrated cells, it appears at ~1225 cm^{-1} [11]. The intensity of the symmetric phosphate stretch is very sensitive to hydration and appears much stronger in hydrated cells and tissues compared to fixed or dehydrated samples. In fact it has been shown that a spectrum of an air-dried enucleated human erythrocyte is almost identical to a spectrum of an air-dried nucleated avian erythrocyte in the phosphodiester region [38]. It is only when the avian erythrocyte is hydrated that one clearly observes the DNA bands and in this state they can actually be accurately quantified in intact simple cells [12]. Figure 15.3 shows a series of transmission spectra recorded of nucleated chicken avian erythrocytes as a function of hydration. The red spectra represent the starting point where the cells are totally dehydrated. Water vapour was gradually introduced into the sample cell and the cells begin to hydrate (green spectra). The blue spectra show the cells in a totally hydrated state. As the cells become hydrated the intensity of the symmetric phosphate stretch dramatically increases and the shift of the asymmetric phosphate stretch from 1225 cm^{-1} to 1244 cm^{-1} is evident. In general, these modes appear very

broad in dehydrated cells and fixed tissues and are often overlapped by glycogen, collagen, phosphorylated proteins and phosphodiester groups associated with lipids and fatty acids. A shoulder feature on the amide I mode appearing at 1712 cm^{-1} in dehydrated cells and at 1716 cm^{-1} in hydrated cells is assigned to in-plane base C=C and C=O double bonds stretching from DNA. A band at 1050 cm^{-1} assigned to the C-O stretching vibration from deoxyribose shows a dramatic increase in intensity as a function of increasing hydration. The band at 967 cm^{-1} is assigned to a C-C stretching vibration based on its similarity to the 950 cm^{-1} band observed diethyl phosphate anion [39]. This band also appears sharper and more intense in the hydrated state.

It is hypothesized that as the DNA transforms from the B-DNA conformation to the A-DNA form the overall molar extinction coefficient decreases because of the disordered alignment of the individual molecular dipoles in the phosphodiester backbone. This has the effect of broadening the asymmetric and symmetric phosphodiester stretching modes to the point where they are overlapped by bands from other molecules. In summary, it is important to control the humidity when working with air-dried cells. If one wants to visualize the DNA more clearly then the cells should be hydrated or a fixative that preserves the B-DNA form needs to be applied. Any type of fixation will result in the formation of A-DNA, broadening the DNA bands to the point where they are not perceptible against the superposition of other bands in the region.

15.3 Instrumentation for FTIR Imaging

15.3.1 FTIR Microscope Configurations

The coupling of an FTIR spectrometer with a Cassegrain objective microscope has introduced a powerful tool for simultaneous analysis of the chemical structure and spatial distribution in heterogeneous materials of micron size (~2.5–25 μm). This has opened up the possibility of diagnostics of biochemical changes to the level of the single cell. IR microscopes are designed to operate in either transmission or reflection mode. In addition, the microscope can be equipped with a micro-ATR accessory (usually germanium crystal) allowing collection of attenuated total reflection IR spectra (ATR FTIR). Transmission or transflection modes are the most popular modes used to investigate cells and tissues. To collect FTIR spectra in transmission mode, the IR beam from the interferometer passes to the Cassegrain condenser—a two mirror optic. There, the beam is focused on a sample placed on an *x, y, z* motion stage and transmitted through the sample to the Cassegrain objective, and next to an adjustable aperture. After that, the radiation is focused onto a detector, which records the spectrum from the sample area limited by the aperture. The region of interest is imaged by using a CCD camera. Figure 15.4 illustrates a simplified scheme of the IR microscope.

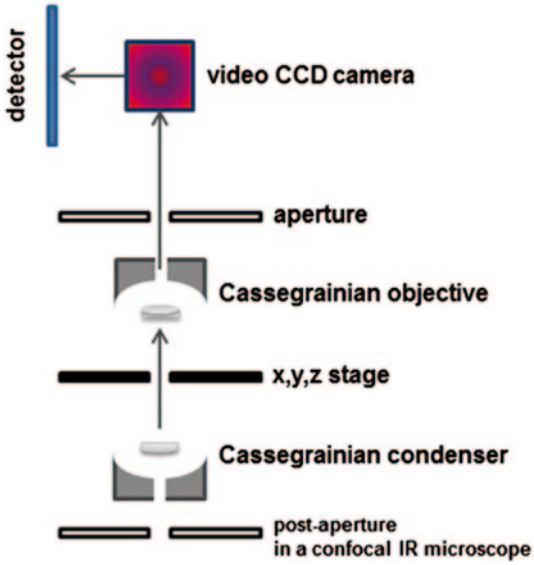


Fig. 15.4 Scheme of an FTIR microscope

For reflection and transflection mode, the Cassegrain system above the *x, y, z* stage operates as both the condenser and objective. The reflected beam is directed to the aperture through a set of mirrors, and then onto the detector. In transflection mode the optical path length and hence the absorbance is doubled because the beam passes twice through the sample. In straight reflection mode the signal is usually weak because fewer IR photons are reflected from relatively non-reflective surfaces like cells or tissues.

In a typical IR microspectrometer the aperture setting determines the sample's area of interest but the spatial resolution of the chemical image is inherently affected by the diffraction limit. The latter depends on the numerical aperture (NA) of the objective and the light wavelength (λ) as expressed in Rayleigh's criterion: 0.61λ/NA. The largest numerical aperture for commercially available Cassegrain objectives is *ca*. 0.65. There are two configurations of IR microscopes, which influence the spatial resolution. In a single-aperture IR microscope, the aperture controls the field of illumination by the IR beam and the spatial resolution is defined by 2λ/3. Hence, for the spectral range from 4000 to 700 cm^{-1}, the diffraction-limited spatial resolution varies from 1.7 to 14.3 μm. Conversely in a confocal arrangement, a second aperture determines the area being analysed and thus improves the spatial resolution (to λ/2) and the image contrast by the reduction in the first and higher-order diffraction rings [40, 41]. The second aperture serves as a pinhole, but less energy reaches the detector resulting in a lower signal-to-noise ratio (S/N). Generally the confocal microscope operates only in conjunction with a synchrotron beamline due to the requirement for a very bright point like IR source. A detailed explanation of the concept and limitations of spatial resolution in IR mapping and imaging has been provided by Carr [40] Lasch [42] and Kazarian [43].

Commercially available infrared microscopes are also designed in two configurations, called infinity correction and non-infinity correction. The infinity corrected design provides flexibility in choice of additional optical elements (objectives, filters), on which the beam is collimated. This system has the focal point at infinity for sharp focus in both visible and infrared images by collimating the two beams within the microscope. The non-infinity corrected microscopes are limited in terms of additional optics, which consequently affects the quality of focus and sharpness of the image. Among manufactures of IR microscopes, Varian (now Agilent) and Thermo Fisher produce the infinity-corrected microscopes, whereas Bruker and PerkinElmer employ the non-infinity corrected design.

FTIR spectra of cells and tissues can be recorded in three modes including transmission, transflection (or transmission/reflection), and ATR. Since the best S/N is found for the transmission mode, this technique is the most common for single cell analysis. For a transmission measurement the sample must be placed on a substrate made of a material permeable for IR light such as non-hygroscopic CaF_2, BaF_2 or CsI windows. To avoid high absorption, the thickness of a sample should be less than 10 μm. These IR substrates create chromatic aberration, which results in a non-linear wavelength-dependent focus increasing in the low-frequency spectral region but this problem is minimized in windows with a thickness less than 1 mm. For tissue measurements Ag/SnO_2 coated infrared reflective slides (Kevley Technologies) are often employed. In this case the measurements are performed in transflection mode, because the beam passes through the sample and reflects off the IR reflective layer. The beam thus passes through the section twice, and consequently a suitable thickness of the tissue samples is usually *ca*. 4 μm. The improved absorbance signal, especially for very thin samples such as tissue sections, the absence of chromatic aberration along with the low cost of the reflective substrates in the comparison to the permeable IR widows make the transflection technique a useful tool for IR microspectroscopy of cell, tissues, and biofilms. However, Filik et al. and Bassan et al. reported that transflection IR spectra can be strongly affected by the electric field standing wave (EFSW) effect [10, 44]. The EFSW within multilayer systems results in an inhomogeneous field due to focusing and interference in both transmission and reflection processes. This leads to the non-linear relationship between absorbance and the sample thickness, which distorts the spectra by raising the absorbance of the amide I and amide II region relative to the high-frequency range above *ca*. 2800 cm^{-1}. Even for differences in thickness of 0.1 μm that are common in morphology of cells and tissues prepared for IR imaging, significant changes in shapes and ratios of IR bands can occur as a result of this effect. Therefore, transflection spectra of biological material should be carefully examined, to confirm that the observed spectral changes result from the biochemical status of the cell or tissue under investigation and not from spurious artefacts generated by the transflection mode of the measurement. This can be overcome to some degree in the case of tissues by having very even tissue sections and by using second derivatives, which are sensitive to shifts in band position and do not solely rely on direct absorbance [10].

Since the numerical aperture NA is defined as $NA = n \sin\theta$, where n is the refractive index of the medium, in which the optics are surrounded and θ is the half-angle

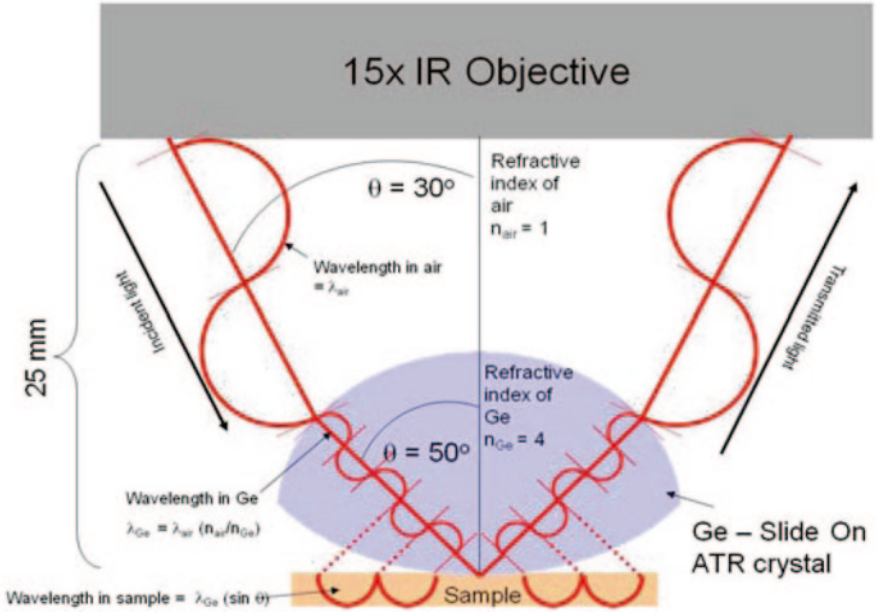


Fig. 15.5 Schematic diagram showing how the IRE element increases the spatial resolution in ATR FTIR microspectroscopy with the use of Ge ATR accessory. Illustration of courtesy of Dr. M. Kansiz and Agilent Technologies

of the light cone that can enter or exit the optical element, the spatial resolution can be improved by making contact with the sample with a material with a high refractive index. This can be achieved using attenuated total reflection (ATR) microspectroscopy with a single-reflection hemispherical internal reflection element (IRE). For example, a germanium (Ge) crystal has a large refractive index (n=4) that simply increases the spatial resolution four times in comparison to transmission and transflection modes, where n for air is equal to 1. In addition, the outer ray angle of incidence (θ) also increases from 30° to 50°, resulting in a further increase in spatial resolution. In such a case, the numerical aperture increases from 0.5 to 3.1. This is illustrated in Fig. 15.5.

Typically, the contact area between the ATR crystal and the sample is *ca*. 100×100 μm^2. Thus, the disadvantage of the ATR FTIR technique is the necessity for good contact between sample and ATR crystal. This contact can distort or damage soft samples of biological origin. The poor quality of contact effects absorbance of the spectral bands, resulting in low absorbance values, which are normally expected to be strong. Moreover, the ATR FTIR spectrum only reflects chemical information from a certain depth of penetration for a given IRE. The depth of penetration can be estimated based on the values of refractive indices of the IRE and the sample, the angle of incidence, and the wavelength of the IR beam [43]. It also can be strongly affected by dispersion for strongly absorbing materials, complicating the determination of the real value. Nevertheless, the depth of penetration is rather small and

it does not exceed 2–3 μm in cells and tissues [43]. Compared to transmission or transflection modes, ATR FTIR spectra do not show artefacts due to Mie scattering and the sample preparation is generally straight forward. Potential applications of ATR FTIR imaging have been reported for both cells and tissues [45–48]. The other advantage of this technique is that normalisation for the sample path length is usually not required because the depth of penetration is constant throughout the medium.

15.3.2 Linear and FPA Detectors

IR microspectrometers operate with the use of a Mercury-Cadmium-Telluride (MCT) detector, which is more sensitive to small amounts of energy delivered through the aperture than the deuterated L-alanine doped triglycine sulphate (DLaTGS) detector typically used for benchtop FTIR spectrometers. A single point MCT detector is a $Hg_{1-x}Cd_xTe$ alloy, in which detection of IR radiation begins with the excitation of an electron from the valence band into the conduction band. The minimum photon energy required is equal to the bandgap, E_g. The bandgap of a MCT device is a function of its composition ratio "x" of CdTe to HgTe, and the temperature of the material [49, 50]. It also affects the cut-off wavenumber. The single and linear MCT detectors employed in IR microscopes operate in the photoconductive mode with the cut-off at *ca*. 650 and 720 cm^{-1}, respectively, whereas a matrix of MCT detectors (MCT FPA, MCT Focal Plane Array) acts as a photodiode and enables recording an IR spectrum from 850 cm^{-1} [49, 50]. The latter belongs to the third-generation MCT devices—hyperspectral array, generating a hypercube image. For such an application, standard second-generation HgCdTe arrays are enhanced with optical coatings and/or proximal filters [50].

Single point MCT detectors, which operate in IR microscopes, usually have a size of 50×50, 100×100, or 250×250 μm^2. A smaller area of the MCT element is usually utilized, since only a part of the detector is illuminated by the infrared beam, and the remaining part only contributes noise. A linear MCT detector consists of a 16×1 MCT array (400 μm × 25 μm). For instance, the PerkinElmer microspectrometer system coupled with a Spectrum Spotlight infrared microscope permits 1:1 or 4:1 imaging, resulting in sample areas of 25×25 or 6.25×6.25 μm^2, respectively, projected onto each detector element. Each element of the linear detector records the spectrum of the corresponding projected pixel size onto the sample. Both the single and linear detectors can operate in conjunction with thermal or synchrotron sources [51, 52]. Chemical images are generated by mapping, i.e. the spectral signal is delivered to the detector in a stepwise manner from different areas of a sample in a grid manner with the use of a chosen aperture size and a computer-controlled motorized *x, y, z* stage. The spatial resolution in this case is also limited by the size of the aperture. Lasch and Naumann have showed that chemical maps of human colon mucosa recorded with the use of a linear detector in 1:4 imaging mode have a spatial resolution of 8.8 μm at 4000 cm^{-1} [42]. Due to the small aperture transmissivity, the point and linear mapping of a small region of interest produces an inefficient S/N

ratio, and consequently requires a large number of scans. This obviously elongates time of data collection.

A MCT FPA detector is built of up to 128×128 elements. Each pixel is 40×40 μm^2, but the pixel projection onto the sample plane is typically 5.5×5.5 μm^2. It is a common misconception that pixel projection defines spatial resolution of an IR image. This value does provide one limit to the achievable spatial resolution. The 128×128 MCT FPA detector allows simultaneous collection of 16,384 spectra from all positions within the sampling area of *ca*. 700×700 μm^2 in one tile (one "snapshot"). This method of data collection is called hyperspectral imaging. This dramatically reduces the time needed for data acquisition. Compared to single element or linear mapping, the advantage of the FPA technique is improved spatial resolution with comparable S/N ratio. In fact, the FPA system is apertureless, so diffraction effects are limited through full illumination of each pixel in the imaging array, contrary to mapping microscopes. This results in recording chemical images close to the diffraction limit along with a very good S/N ratio. This is particularly important for recording spectra of cells. Lasch and co-workers have demonstrated that the use of the mapping and imaging techniques affect data collection time, but does not impact on the spectral features of oral mucosa cells, showing a similar detection limit for both techniques [27]. The utility of each technique for specific cases and their respective limitations are illustrated in detail in [53]. Modern IR microspectrometers also enable hyperspectral imaging in ATR mode. This technique has been patented by Agilent. As mentioned above, the use of the germanium IRE provides a 4-fold increase in spatial resolution and sampling area of approx. 70×70 μm^2, since pixel projection is 1.1×1.1 μm^2. However, Kazarian and Chan have reported that spatial resolution of micro-ATR imaging is affected by non-uniform illumination of the Cassegrain objective, resulting in a non-uniform numerical aperture along the different axes of the image plane [43]. They determined that when applying micro-ATR imaging to a polyurethane-PMMA interface, the spatial resolution achieved at 1600 cm^{-1} with a vertically oriented sample was 4 μm, whereas when the sample interface was oriented horizontally a 50% reduction in spatial resolution was obtained compared to the vertical orientation.

15.3.3 Synchrotron Configuration

Conventional FTIR spectrometers are equipped with a black body source (globar) that gives a much lower density of infrared radiation compared to a synchrotron-based IR source. In fact, an electron storage ring in a synchrotron generates a highly collimated beam of IR photons produced from a point source, which exhibits a brilliance of 100–1000 times greater than that of a globar [54, 55]. Such radiation emitted into a narrow range of angles can be delivered through a very small aperture (down to 3 μm^2) to the area of interest, on which it becomes diffraction limited. But the S/N ratio is significantly improved. Since, a globar source illuminates a *ca*. 100 μm^2 area, the aperture settings are in the range of 20–100 μm^2 for

single-point mapping. As discussed above, the spatial resolution is improved in an imaging system by incorporating a MCT linear array, but the smallest aperture is set to 10 μm^2 for practical reasons. Miller and Smith have demonstrated that the brightness of a synchrotron source is almost 10 times greater than that of a globar for a 10 × 10 μm^2 aperture, but it does not differ significantly for aperture sizes larger than 60 × 60 μm^2 [51]. Moreover, IR microscopes coupled with synchrotron-based sources exhibit higher spatial resolution compared to bench top instruments due to their confocal mode, i.e. by using two passes of light through the same pinhole [40]. The magnification of a Cassegrain objective is often 36 × and 74 × with NA of 0.5 and 0.65, respectively. This allows for the acquisition of spectra with an aperture down to *ca*. 3×3 μm^2 and S/N ratio *ca*. 1000 times greater compared to a globar source [56].The use of the largest commercially available numerical aperture (0.65) enables one to achieve a spatial resolution of *ca*. 0.6 μm for the shortest wavelength at 2.5 μm [57]. In addition, the chemical image contrast is effectively improved in the synchrotron experiment that resolves substructures within a very small sampling area. Other advantages of the synchrotron light are its pulse time nature (10–100 s of picoseconds) and a high degree of polarization, which are dependent on the beam line construction [51].

The confocal configuration for an IR microscope is an option that can only be incorporated for an FTIR spectrometer coupled to a synchrotron light source. In addition, almost all IR synchrotron facilities around the world operate in the single-point mapping mode, considerably lengthening the data acquisition time. For instance, acquiring fully diffraction-limited maps of a 30 × 30 μm^2 area takes 2–4 h [58].Thus, the coupling of a bright synchrotron IR beam with an apertureless FPA detector to maximize spatial resolution and achieve fast hyperspectral imaging has become a need. Such a design has been theoretically proposed by Carr et al. [59] and tested by Moss, Petibois, and Marcelli [60, 61]. However, these experiments did not overcome a fundamental problem of inhomogeneous illumination of all the pixels on the FPA detector by the synchrotron beam, hence a synchrotron IR source typically fills a 10–20 μm^2 area. On the other hand, these experiments have shown that spatial resolution near 2.3 μm for light wavelength at 2–4 μm is achievable and spectra collected from the illuminated area of an FPA detector showed a higher S/N ratio compared to thermal sources [60, 61]. A comprehensive approach for the use of a synchrotron beam with an FPA detector has been recently published by Hirchmugl, Bhargava, and Gough [57, 62]. A new design for a synchrotron-based IR imaging system was implemented at the Synchrotron Radiation Center in Stoughton Radiation, Wincosin (mid-infrared beamline IRENI). They have constructed an infrared imaging system designed and optimized to illuminate a large field of view, and which achieves a true diffraction-limited spatial resolution. Their approach was to couple multiple low-emittance synchrotron beams with a large FPA detector [57]. Radiation from a bending magnet is split into 12 beams by 12 toroidal mirrors,

redirected by plane mirrors, collimated by paraboloidal mirrors, and then rearranged by another set of plane mirrors to illuminate a sampling area [62]. The 3×4 beam matrix homogenously illuminates an area up to 60 × 40 μm^2 (96 × 96 pixels) with a pixel projection of 0.54 × 0.54 μm^2 on the sample plane with the use of a 74×, 0.65 NA objective. According to the Rayleigh criterion, for this type of objective, diffraction limited resolution can be achieved with pixel size not larger than 0.59 μm. The effective geometric pixel area of IRENI meets this criterion [62]. Since, this pixel size is smaller than the maximum pixel size allowed for correct spatial sampling, diffraction-limited images are possible even at the wavelength of 2.5 μm. Furthermore, such a design significantly reduces acquisition time from 13 days for raster scanning of a 280 × 320 μm^2 area using a synchrotron-based IR microscope at diffraction-limited resolution to *ca*. 30 min for FPA imaging. The most important advantage of the IRENI setup is achieving high-contrast images, which clearly resolve subcellular structures of 2–7 μm diameter in a comparable manner to H&E stained images [57, 62]. Spectral information derived for different substructures is improved due to pixel localization. In addition, measurements of 1 μm polystyrene beads have shown that the synchrotron-based FPA system can reach a spectral limit of detection of 6 ± 1 fmol in a single 0.54 × 0.54 μm^2 pixel area [57]. While there is certainly an improvement in contrast and spatial resolution the actual signal-to-noise ratio of the resulting spectra are somewhat compromised because in order to illuminate the FPA evenly the lower Cassegrain condenser must be defocused over the entire array. Defocusing the condenser results in poorer quality spectra and so this is something that should be addressed in future configurations.

15.4 Spectral/Image Processing Techniques

Prior to image reconstruction the spectra are usually subjected to some form of pre-processing to compensate for scaling differences. Scaling differences arise from path length effects, scattering effects, source or detector variations, or other general instrumental sensitivity effects. There are two general types of FTIR spectra pre-processing methods namely filtering and model-based methods. The filtering methods, include polynomial or linear baseline correction, calculating second derivatives, smoothing to remove noise and normalization for differences in sample thickness. Model-based methods enable the quantification and separating out of the different types of physical and chemical contributions to the spectra. The information in this case is not lost and statistical estimates of the mathematical parameters involved in the filtering can be obtained [63]. The type of physical contributions to the spectra will dictate the pre-processing regime that is required.

15.4.1 Data Pre-Processing Steps

15.4.1.1 Baseline Correction and Derivative Analysis

Baseline aberrations in FTIR image data can result from Mie scattering due to particles that are in the range of infrared wavelengths typically between 2.5 μm and 25 μm or from light dispersion from tissue boundaries where there is a large change in the refractive index compared to the background measurement. Other baseline features can arise due to detector cut off and substrate effects. Environmental effects like heating can also cause baseline offsets. If the baseline effect were simply caused by an offset due to heating of the sample then a linear baseline correction would suffice. If the base line exhibits a gentle curve then a polynomial approach may be required [64]. These approaches work well when quantification based on integrated area under a peak or a ratio of two bands would suffice in the analysis.

For data that will be used in modelling or a search database it is often better to perform a derivative to remove the baseline contribution. This has the added advantage of better resolving the absorption bands, which are often broad and overlapped. First derivatives emphasize bandwidths, because the first derivatives of most band shapes have maxima or minima near 75 % of the original peak height [65]. To this aim, the second derivative analysis is widely applied to the measured spectra as it highlights band positions and separations. In a second derivative spectrum, the overlapping absorption components in the spectrum are displayed as negative bands. This analysis requires spectra with high signal-to-noise ratio and free of vapour absorption, as second derivative band intensity is inversely proportional to the square of the original band half-width, this leads to an enhancement of the relative contribution of sharp lines due to noise and vapour.

The theory is based on the original work by Butler and Hopkins in 1970 [66] and further modified by Susi and Byler [67]. The following description and notation is based on the theory and notation proposed by Susi and Byler [67].

FTIR spectra are digitally encoded with one data point n every $\Delta k/(2^m) = \Delta W$ frequency units, where Δk is the nominal instrumental resolution selected in cm^{-1}, m is the number of times the interferogram is zero-filled before the Fourier transform is performed. To calculate the second derivative:

$$A''_n = \left(A_{n+1} - 2A_n + A_{n-1}\right) / \left(\Delta W\right)^2 \tag{15.1}$$

where A_n is the absorbance at data point n of the original spectrum.

In this case the second derivative is not smoothed. To smooth the spectrum a Savitsky-Golay function is usually convolved to the derivative function [68]. Depending on the level of noise the smoothing may typically require 9, 11 or 13 points.

If the intrinsic line shape of a vibrating oscillator is approximated by a Lorentzian function in the form:

$$A = \frac{\sigma}{\pi\left(\sigma^2 + k'^2\right)} \tag{15.2}$$

$$= \frac{1}{\sigma\pi\left(1 + Bk'^2\right)} \tag{15.3}$$

where A is the absorbance, 2σ is the width at half height, $k' = (k - k_0)$ is the frequency with respect to the band centre at k_0, and $B = 1/\sigma^2$. Substituting Eq. 3 into Eq. 1 gives the second derivative.

$$A'' = \frac{-2B\left(1 - 3Bk'^2\right)}{\sigma\pi\left(1 + Bk'^2\right)^3} \tag{15.4}$$

The band position for the second derivative is identical with the wavenumber value of the original band centre k_0. The half width of the second derivative, σ^{II} is related to the half width of the original line by

$$\sigma^{II} = \sigma / 2.7 \tag{15.5}$$

while the peak intensity A_0'' of the second derivative is related by the following expression

$$A_0'' = -2A_0 / \sigma^2 \tag{15.6}$$

Thus the peak height of the second derivative is proportional to the original peak height and inversely proportional to the square of the original half-width. Of course for real spectra, where there are overlapping bands, which deviate from a true Lorentzian the situation is more complex but nonetheless these formulas provide a reasonable approximation for second derivative interpretation [67].

In summary the second derivative corrects baseline shift, resolves overlapping bands, reduces replicate variability and amplifies spectra variations without having to model for spectral variation.

15.4.1.2 Normalization

Normalization is often employed to compensate for the differences in sample quantity (concentration and/or path length) in a spectral database or image hypercube. Normalization pre-processing methods attempt to correct for these effects by identifying some aspect of each sample, which should be essentially constant from one sample to the next, and correcting the scaling of all variables based on this characteristic. Typically, normalization should be performed before any centering or scaling or other column-wise pre-processing steps and after baseline or offset removal. The presence of baseline or offsets can impede correction of multiplicative effects. The effect of normalization prior to background removal will, in these cases, not improve model results and may deteriorate model performance.

Prior to classification, a normalization or standardization step is often required so that the extracted spectral information can be used for comparison between classes. One type of normalization is known as min/max normalization and can be applied providing the peak absorbance band is constant between samples, which is often not the case. A number of peaks in a biological sample are often used as maximum peaks for min/max normalization and providing the chosen peak is stable and consistent then this method is feasible. Such bands include the amide I, amide II and the CH_2 asymmetric stretching peak. In this type of normalization the peak height of the chosen band is adjusted to 1 by multiplying the whole spectrum by a multiplication factor. In this case the absorbance information in the spectrum is lost and thus the normalized spectra are relative absorbance spectra and only suitable for qualitative analysis. In most studies min/max normalized spectra are only used for display purposes, whereas for classification purposes non-normalized spectra are normally incorporated into the model. An alternative to min/max normalization is vector normalization where the normalization is not based on a single peak but rather each spectrum is divided by the Euclidean norm. Vector normalization is often applied to second derivatives prior to cluster analysis, because second derivatives have values below zero whereas min-max normalization can only process FTIR spectra with positive values. As a general approach for tissue analysis and providing the baselines are not affected by Mie scattering or light dispersion and where the signal-to-noise is excellent we would take the second derivative and then vector normalize the data prior to cluster analysis.

15.4.2 Light Scattering Correction

Normalization and derivative analysis can correct for differences in sample thickness and sloping baselines but if light scattering and dispersion distort the spectra then a modelling approach is required. Scattering results when the light is deviated from the detector resulting in broad oscillating baselines with shifts in peak position and band shape distortion. One of the main forms of scattering observed in the FTIR spectra of cells is Mie scattering. Mie scattering is manifest when the scattering particles are in the range of infrared wavelengths (5–12 μm), which is the size of small cells and nuclei.

The Mie scattering efficiency can be determined using the van de Hulst approximation equation [69]:

$$Q = 2 - \frac{4}{\rho}\sin\rho + \frac{4}{\rho^2}(1 - \cos\rho) \tag{15.7}$$

$$\rho = \frac{2\pi d(n-1)}{\lambda} \tag{15.8}$$

where d is the diameter of the particle, n is the real refractive index and λ is wavelength.

15.4.3 Multiplicative Signal Correction

The first types of approaches applied to correct spectra using modelling methods were Multiplicative Signal Correction also termed Multivariate Scatter Correction (MSC) [70]. MSC is a chemometric method that models for the effect of light scattering. The method was originally developed for applications of near-IR spectroscopy in food science [71]. In MSC, each spectrum is regressed in a set of related samples. In MSC a mean (or reference spectrum) is used to estimate the intercept and slope of an estimated regression equation [72]. Subtracting the intercept and dividing by the slope corrects each spectrum. The MSC will only work providing there are negligible differences between the reference and corrected spectrum or alternatively if the regression analysis is performed on a region devoid of chemical information and thus only influenced by multiplicative light scattering [72].

The basic equation for a typical MSC follows a simple linear regression:

$$\boldsymbol{Z}_{raw} = a + b\boldsymbol{Z}_{corr} \tag{15.9}$$

where $\boldsymbol{Z}_{raw}$ is the unprocessed spectrum prior to correction, a is the constant for offset baseline correction, b is the normalization constant, and $\boldsymbol{Z}_{corr}$ is the corrected spectrum.

Thus, the MSC corrects for baseline offset and normalization and does not take into account non-linear effects such as Mie scattering and dispersion.

15.4.4 Extended Multiplicative Signal Correction

Extended Multiplicative Signal (or Scatter) Correction (EMSC) is an extension of MSC based on linear statistical regression modelling that can model for non-linear effects such as Mie scattering by utilizing a multivariate bilinear model [73]. The EMSC theory was originally proposed by Martens and Stark for correcting light scattering in near-IR spectra of solids, slurries and optically dense liquids [74]. An example of a multiplicative effect that can be modelled using the standard EMSC is light scattering variation in response to changes in the optical path length [74].

In this case, the spectrum can be described as the superposition of the pure absorbance spectrum with both baseline offset and sloping baselines removed and can be written as:

$$\boldsymbol{Z}_{raw} = a + (m\bar{\nu}) + (b\boldsymbol{Z}_{ref}) + \boldsymbol{E} \tag{15.10}$$

Where m is the gradient of the sloping baseline, $\boldsymbol{Z}_{ref}$ is the reference spectrum, which can be an averaged spectrum, and $\boldsymbol{E}$ is the un-modelled residual variance i.e. the interesting variations in the chemical information within the spectrum, which are not accounted for by the reference spectrum. The corrected spectrum can be calculated after the constants *a, m* and *b* are estimated using a least squares approach.

$$\boldsymbol{Z}_{corr} = \frac{\boldsymbol{Z}_{raw} - \left(a + m\bar{\nu}\right)}{b} \tag{15.11}$$

Kohler et al. [73, 75] incorporated Mie scattering correction into EMSC (Mie-EMSC) by using principal component analysis (PCA) to decompose ~150 artificially generated Mie curves calculated for a range of Mie scattering particles and a range of refractive indices [73]. However, if the scattering particles exhibit absorptions then the wavelength dependent refractive index needs to be taken into consideration.

A further extension of this algorithm was championed by Bassan and Gardner and named *Resonant* Mie Scatter EMSC (RMieS-EMSC) [10, 76–81]. Resonant Mie scattering theory recognises that real samples are absorbing, and hence a full description of the Mie scattering must account for the fact that the Mie scattering efficiency is dependent upon the refractive index, which changes rapidly on passing through an absorption resonance. Bassan and Gardner modified the Mie-EMSC algorithm originally developed by Kohler to correct the distorted spectra by fitting a refractive index spectrum obtained using the Kramers-Krönig transform of the absorption spectrum. The following description is a condensed version of that described by Bassan and Gardner [77]. RMieS-EMSC follows the Mie-EMSC approach in that a set of Mie scattering efficiency (*Q)* curves are generated by selecting a range of particle diameter sizes (d) and a range for the real refractive index (n). To generate the required complex refractive index parameters a Kramers-Krönig transform is performed to obtain the real refractive index spectrum, $n(\bar{\nu})$:

$$n_{kk} = n\left(\bar{\nu}\right) - \langle n \rangle = \frac{2}{\pi}\mathbf{P}\int_0^{\infty} \frac{s \times k\left(\bar{\nu}\right)}{s^2 - \bar{\nu}}\, ds \tag{15.12}$$

where $k(\bar{\nu})$ is the imaginary part of the refractive index, $\langle n \rangle$ is the average refractive index, n_{kk} is the real refractive index spectrum minus the average real refractive index, and **P** is the Cauchy Principal of improper integrals. The key to usefully performing the Kramers-Krönig transform is to recognise that the imaginary refractive index (the k term) can be replaced by the reference absorption spectrum $\boldsymbol{Z}_{ref}$ as they are very nearly proportional.

The set of Q curves can then be placed into a matrix **Q** with each row representing a Q curve. Using PCA, the matrix can be decomposed into a small number of principal components (PCs).:

$$\mathbf{Q} = \mathbf{T}\mathbf{P}^{\mathbf{T}} \tag{15.13}$$

where **T** represents the score matrix and **P** the loadings matrix. The first few loadings describe the majority of the variance. These loadings are incorporated into the EMSC model in the following manner.:

$$\boldsymbol{Z}_{raw} = a + m\left(\bar{\nu}\right) + b\boldsymbol{Z}_{ref} + \sum\nolimits_{1}^{7} g_i \boldsymbol{p}_i + \boldsymbol{E} \tag{15.14}$$

where the g_i are the weights applied to the loading vector $\boldsymbol{p}_i$. This expression can be re-arranged to give the corrected spectrum.:

$$\boldsymbol{Z}_{corr} = \boldsymbol{Z}_{raw} - \left(\frac{a + m\bar{\nu} + \sum\nolimits_{1}^{7} g_i \boldsymbol{p}_i}{b} \right) \tag{15.15}$$

15.4.5 Other Methods to Correct for Dispersion and Mie Scattering

The Diem group introduced a phase correction approach that is based on the notion that the complex Fourier transform can be separated into the real and imaginary parts of the spectrum or interferogram by varying the phase angle between them. Using the Mertz algorithm, a wavenumber-dependent phase angle can be calculated and phase corrected spectra computed by a complex forward Fourier transformation of the real and imaginary part of the spectrum or interferogram [82]. The advantage of this approach is the computation time is dramatically reduced.

15.4.6 Unsupervised Hierarchical Cluster Analysis

There are three main ways to present FTIR spectral data, namely univariate (chemimages), multivariate and pattern recognition. Univariate methods are typically based on determining the absolute intensity of a band, or the ratio of two bands or are based on determining the integrated area of a band between two baseline points. These maps provide quantitative information on macromolecules in a cell or tissue sample but are generally not diagnostically useful. Multivariate methods make use of spectral similarity to generate false-colour maps. There are many approaches to generate multivariate images from spectral data hypercubes. These include Unsupervised Hierarchical Cluster Analysis (UHCA), K-nearest neighbours (K-means), Principal Component Analysis (PCA) and artificial neural networks (ANNs) to name a few. Unsupervised Hierarchical Cluster Analysis (UHCA) is a proven technique for image reconstruction with FTIR data [83–85]. In cluster analysis, a matrix is calculated that expresses the similarity, or ‘‘distance’’, between each spectrum

and all other spectra in the data set. This distance is usually calculated using Person's Correlation Coefficient as a measure of similarity. The complimentary distance measure is known as D–values but other distance equations including City block, Euclidean and normalized Euclidean can also be applied. The Pearson's correlation approach minimizes spectral heterogeneity because dividing through by the residual standard deviation normalizes the distance.:

$$r_{jk} = \frac{\left(\sum_{i=1}^{n} x_{ji} x_{ki}\right) - n\bar{x}_j \bar{x}_k}{\sqrt{\left(\sum_{i=1}^{n} x_{ji}^2 - n\bar{x}_j^2\right)\left(\sum_{i=1}^{n} x_{ki}^2 - n\bar{x}_k^2\right)}} \tag{15.16}$$

Here r_{jk} is the correlation coefficient, x_{ji} represents the spectrum, x_{ki} represents the spectrum to be compared and n is the total number of the spectra.

For two very similar spectra the $r_{jk} \approx 1$ (i.e. distance ≈ 0), whereas for completely dissimilar spectra this value approaches 0 (i.e. distance is large). Subsequently, the two most similar spectra in the hypercube are merged into a ''cluster'', and a new distance matrix column is calculated for the new cluster and all existing spectra. The process of merging spectra or clusters into new clusters is repeated, and the distance matrix is recalculated until all spectra have been combined into a few clusters. This process combines the most similar spectra into the same cluster, while keeping track of which spectra have been incorporated into each cluster. The clustering or merging of spectra can be achieved using a variety of equations but the one often used is Ward's algorithm [86].

Assigning a colour to each spectral cluster, and then displaying this colour at the coordinates at which each spectrum was collected generates a pseudo-colour map. The mean spectra can then be extracted for all clusters and used for the interpretation of the chemical or biochemical differences between clusters. The cluster number can be adjusted so that a good correlation with the pathological images can be obtained.

15.4.7 Neural Networks

Artificial neural networks (ANN) (an artificial intelligence system for machine learning) model for non-linear relationships in the data hypercube and can be trained to learn histological and pathological features. Once trained, neural networks have the advantage of high speed compared to other multivariate techniques including UHCA.

Many networks make use of the backpropagation algorithm for training the network and utilize multiple layers of weighted neuron units each defined by an activation function (c.f. Fig. 15.6). The backpropagation algorithm is a supervised learning rule, which requires a training set of inputs with known-to-be-correct target

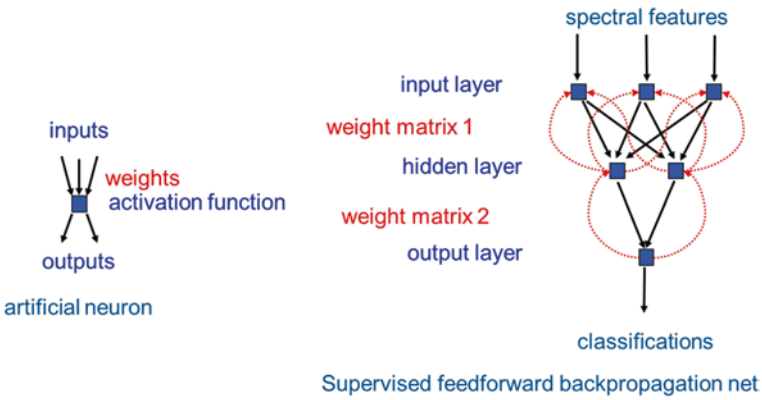


Fig. 15.6 A multilayer artificial neural network system for automatic classification of spectra

output classifiers. During each cycle of the learning phase, the training input activations are first passed forward to compute output activations (i.e. the classifications). Next, the output classifications are compared with the training set target classifications, and based on the errors a backpropagation process is deployed through the network to adjust each of the neuron's input weights. At some predefined point of minimal errors, the learning cycle completes (it is common practice to use a cross-validation approach against a test set of spectra when determining the remaining error to guard against the possibility of obtaining an overtrained network). It is necessary to test the final trained neural network against a fully independent validation set of spectra, again with known target output classes, to validate the generalization performance of the trained network. It was shown by Vapnik and Chervonenkis [87] that given a finite sized set of training data there is an upper limit to the number of neuron interconnections that can be included if the network is to achieve good generalization [88]. In practice, for FTIR analysis, this limits the number of spectroscopically derived values that can be included as inputs for the ANN. One approach to reducing the dimension of the input data is to first perform a feature selection on the spectra. For example, only absorbance band features of chemical interest or regions of the spectra containing the most variance within the data set might be presented as inputs to the network.

One advantage of artificial neural networks for automatic classification (i.e. diagnosis) of spectra of tissues and cells is their potential to discover non-linear separation of classes. To enable this feature a non-linear sigmoid activation function is employed to determine if the stimulation received at a neuron's inputs is sufficient to generate output activation for feed forward to other neurons. A disadvantage of ANN analysis is that it is difficult to interpret the model to determine which of the inputs (i.e. which spectral information) are most important for successful class prediction.

15.5 Tissue Analysis

The remainder of this chapter will concentrate on some of the many spectroscopic studies undertaken from the 1990's onwards, associated with medical diagnostics of human diseases such as cervical cancer, neurodegenerative pathologies, and dysfunctions of cardiac and liver tissues. They will include investigations on animal models as well as tissue samples collected from patients. These studies are mainly based on searching for "infrared biomarkers" within the complex mid-infrared spectrum that are indicative of a pathological, physiological or a biological state (e.g. different cell types within a tissue), which can be used to discriminate statistically classes of samples. Very often, this statistical approach is accompanied by attempts of understanding the biochemical information delivered by IR spectra and their correlation with the existing diagnostic methods. Animal models usually serve to mimic the progression of diseases or to assess therapeutic effects of the treatment. Investigations with the use of FTIR imaging have evidenced a powerful ability of the technique to be a complementary tool in monitoring and understanding mechanisms of the pathogenesis and the etiology of disease. This is actually the key to further development of clinical applications of FTIR spectroscopy to disease detection.

15.5.1 Cervical Tissue

Worldwide, cervical cancer is one of the most prevalent forms of cancer in females, especially in a form of invasive squamous cell carcinoma. This disease is screened cytologically using the Papanicolaou test (Pap smear), while confirmation of the diagnosis of cervical cancer or pre-cancer usually requires a biopsy of the cervix. HPV DNA has been detected in 99.7% of cervical cancer cases thus confirming HPV as an etiological factor for the disease [89]. HPV types have been classified into 5 genera (alpha, beta, gamma, mu and nu) and 33 species. The alpha genus is clinically the most important and many of these types infect the mucosal epithelium. Of these, a subset of 10–15 HPV types have been associated with cancerous lesions [90]. HPV-16 and HPV-18 are found in 50–65% and 7–20% of cases of cervical cancer, respectively [90].

The "gold standard" for diagnosis of cervical cancer is the staining and pathological analysis of a cone biopsy. The diagnosis of the cone biopsy is classified by the three-tier cervical intraepithelial neoplasia (CIN I, II and III) system. CIN, a precursor to cervical dysplasia or carcinoma in situ (CIS), is diagnosed from biopsies. CIN is classified as mild (CIN1), moderate (CIN2), and severe (CIN3) dysplasia. Progression beyond CIN III is known as Carcinoma in situ (CIS). Since, CIN2 and CIN3 have the highest risk of progressing to CIS, more recently, these groups have been combined into CIN2/3. Virology techniques including immunohistochemistry and electron microscopy, lack the sensitivity and specificity for the routine detection of HPV [90]. On the other hand, serological assays for the detection of anti-HPV antibodies have limited accuracy and consequently their role in the clinical

environment is under review. At present all commercial HPV tests in diagnostic use rely on the detection of HPV nucleic acids in clinical specimens.

The first report highlighting the potential of FT-IR microspectroscopy as a promising tool for screening for cervical malignancies was provided by Wong et al., who have showed reproducible spectral differences between normal and malignant cervical epithelial cells [91]. A similar classification has been shown by using PCA decomposition of IR spectra from 272 patients with normal cervical cells and those with varying degrees of dysplasia and HPV-infected cells [92]. This discrimination reveals characteristic spectral features of the cervical cells. Namely, spectra of the normal tissue show high absorbance of the glycogen band and high $\nu_s(PO_2^-)/\nu_{as}(PO_2^-)$ ratio on the contrary to smear classified as abnormal epithelial cells. In the latter, the $\nu_{as}(PO_2^-)$ band becomes more intense and broader due to increasing content of nucleic acids. Changes in intensity of bands in the high-wavenumber region was observed when applying pressure-tuning FTIR indicating the disorder of methylene chains in membrane lipids of cervical neoplastic cells [92]. The PCA classification was supported by histological results (Pap smear and biopsy), showing 91 and 87% correlation for normal and cancer diagnosis, respectively [92]. Another IR biomarker has been proposed by Mordechai et al., who have studied formalin-fixed tissue from 20 patients with diagnosed histopathologically invasive cervical carcinoma [93]. The RNA/DNA ratios calculated for the 1121 and 1020 cm^{-1} bands are higher in malignant areas than in normal areas of the tissue, possibly indicating higher transcription during carcinogenesis. By using this parameter in conjunction with identified earlier changes in band intensity of glycogen and phosphate groups, two distinct clusters normal and cancerous cervical tissues were reported [93]. However, further studies have demonstrated that the observed spectral changes might not be strictly correlated with the presence of dysplastic cells, but can result from differentiation, maturation, a stage of the menstrual cycle, contamination of mucin, blood cells, and other factors [94–98]. Consequently, these changes can also cause non-specific depletion of glycogen and variation of phosphate group bands. FTIR mapping [99, 100] and imaging [101] of the cervical tissue have given a new insight into overall biochemical characterisation of histopathologically normal and cancerous tissues delivered from biopsies of the same patients. The FTIR mapping of normal and squamous CIS cervical tissues allowed for their discrimination within the basal layer which is important as this region is where cancer cells often originate [99]. The basal layer of the both normal and CIS biopsies was identified by the presence of sharp bands at 1204 and 1283 cm^{-1}. Apart from those bands, a high-intensity band at about 1155 cm^{-1} was found in the normal basal tissue, whereas a broad band near 1240 cm^{-1} became more prominent in the precancerous tissue. Thus, classification of normal and cancerous cervical tissues relied on the ratio of the 1155 and 1240 cm^{-1} bands. This ratio was 1.40 ± 0.91 and 0.29 ± 0.19 for the normal and CIS biopsies, respectively [99]. In turn, Wood et al. have demonstrated spectral characteristics of cervical sections classified to mild and severe dysplasia by using FTIR mapping of 52 different regions of tissue from 10 different patients [100]. The examination of such a huge data bank was performed with the use of multivariate Unsupervised Hierarchical Cluster method (UHCA).

This algorithm provides spectral features of normal cervical transformation zone, associated with anatomical features observed in H&E staining such as the ectocervical and endocervical epithelium, the squamous junction, and red blood cells. The squamous layer shows high concentration of glycogen, whereas the epithelium tissue exhibits the presence of glycoproteins. A good demarcation pattern of squamous epithelium layers was also found for a patient diagnosed with low-grade squamous intraepithelial lesions [100]. Here, most of the differentiation was due to variations in protein bands. Furthermore, the superficial layer exhibits a high glycogen content in contrast with the parabasal layer, in which a reduction in the amide I/amide II ratio is also observed. As expected, IR spectra of the connective tissue show strong collagen bands and no contribution from glycogen. The unexpected low glycogen content was found within the parabasal and superficial layer, which was proposed as an indicator of parakeratosis. Below this layer, an identified cluster was associated with inflammation. Furthermore, this approach clearly detects abnormal areas within a tissue section diagnosed as CINII/III. The spectrum representative of such a metastasis highlights a significant decrease in amide II/amide I intensity ratio, suggesting that the amide II band can be a possible marker of diseased tissue. However, one cannot rule out the possibility that the excellent demarcation of cancer *versus* normal tissue in these sections was the result of light dispersion because of the different refractive indices observed between the highly nuclei dense regions of CIN III dysplasia compared to the normal tissue with pyknotic nuclei and large amounts of cytoplasm. The extracted spectra from these maps did show a significant shift in the amide I region indicating evidence of dispersion. Given that the best demarcation was achieved using the amideI/amide II region and not the phosphodiester region (1200–1000 cm^{-1}) and at that point in time RMie-EMSC had not been developed it is reasonable to assume that the demarcation was based on dispersion and not differences in protein expression. Other spectral features of this abnormal tissue include an increase in intensity of both phosphodiester bands [$\nu_{as}(PO_2^-)$ and $\nu_s(PO_2^-)$]. Figure 15.7 shows H&E stained sections (overview and zoomed in area) of a CIN III diagnosed cervical tissue clearly highlighting regions of cancerous tissue in the darker purple colour. Below are the corresponding FTIR maps after performing Unsupervised Hierarchical Cluster Analysis (UHCA) on the amide I and amide II region (1700–1500 cm^{-1}). The UHCA maps clearly delineate the major histological and histopathological features. The brown is associated with the abnormal cells, light blue lymphocyte exudates, dark blue connective tissue, orange red blood cells and the red are regions are areas of weak absorbance. As shown in the bottom right panel this approach also resolves micro metastases and surrounding lymphocytes. This work has clearly demonstrated that combination of FTIR imaging with unsupervised multivariate image analysis provides a "blind" detection of biological variability within normal and diseased state of the cervical tissue [100] albeit based on physico-chemical differences and not solely chemical differences.

Steller and co-workers performed a similar study on a sample from a patient with a squamous cell carcinoma of the uterine cervix who underwent a radical hysterectomy [101]. 122 chemical images of the tissue were recorded using an FTIR microspectrometer coupled with a 64×64 FPA detector. Spectral processing was

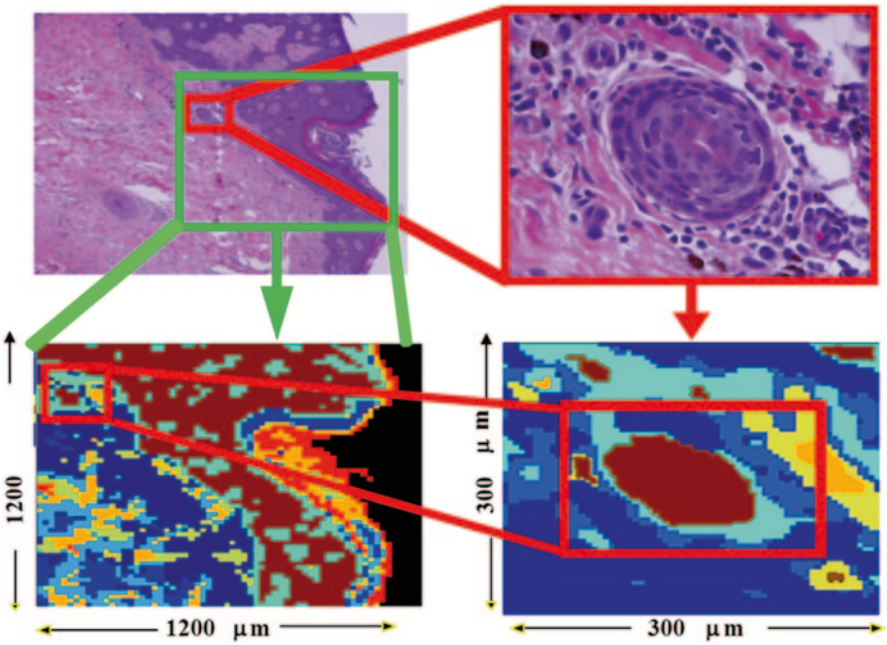


Fig. 15.7 H&E stained sections (overview and zoomed in area) of a CIN III diagnosed cervical tissue along with the corresponding FTIR maps after performing Unsupervised Hierarchical Cluster Analysis (UHCA) on the Amide I and Amide II region (1700–1500 cm^{-1})

carried out by using fuzzy C-means clustering followed by UHCA. IR images were collected from the five major tissue types: cervical stroma, squamous epithelium, squamous cell carcinoma, dysplasia, and peritumoral inflammatory response. Multivariate analysis in the region of 950–1480 cm^{-1} discriminated all tissue layers with some overlap. The cervical stroma and the area of inflammatory response along with the transition layer between them contain a similar set of bands assigned to proteins as the major constituent of these tissues. The variation in intensity of a band at 1082 cm^{-1} was found in spectra of all three tissue types and is a possible indicator of inflammatory response. A higher intensity of the collagen band (a triplet centered around 1232 cm^{-1}) in the cervical stroma is found to be consistent with a lower intensity of this band in the spectra of epithelium, while a cell layer between the basal to parabasal tissues is characterized by a slightly increased absorbance near 1028 cm^{-1}. Despite the clear discrimination of the main components of the cervical tissue, the applied multivariate method did not categorize the basal cell layer, dysplasia, and squamous cell carcinoma. Only clustering in the spectral region of 1420–1480 cm^{-1} showed some spectral discrimination of pathological states. This is associated with shifting of a band assigned to the deformation vibration of the CH_2 and CH_3 groups from 1460 to 1452 cm^{-1}, along with the decrease in its intensity, which is maximum for the basal cell layer, medium for the dysplastic tissue, and minimum for the squamous cell carcinoma [101].

The validity of FTIR microspectroscopy for classifying biopsies with symptoms of cervical intraepithelial neoplasia in different grading (CIN I-III) has been evaluated by Mordechai and co-workers by using probabilistic neural networks (PNNs) and histopathology [102]. The results show the high level of accuracy in discrimination of the grading of neoplasia (CINI, CINII, and CINIII) comparable to a histological method. The false positive identification of the normal as CINI, CINII, and CINIII patients was 9.04, 0.01, and 0.01 %, respectively. The small errors were encouraging for further advanced studies on spectroscopic diagnostics [102]. The same research group has reported that spectral features of a control group and a group of patients diagnosed with cervical cancer result from the decrease in the glucose/phosphate ratio (by 23–49 %) and the increase in the RNA/DNA ratio (by 38–150 %) in carcinogenic compared with normal tissue [103]. In this case, probabilistic neural networks and statistical tests predicted the presence of cancer with a high degree of accuracy (>95 %).

In another experiment, Bogomolny et al. [104] have studied the application of FTIR microspectroscopy for monitoring malignancy progression in murine fibroblast (NIH/3T3) and mouse embryonic (MEF) cells infected with murine sarcoma virus (MuSV). Human cervical tissues at different stages of neoplasia were used for testing. They proposed an array of four biomarkers, which discriminated different premalignant stages with a classification accuracy of 89.5–97.4 %. The identified biomarkers are associated with alterations in the lipids/proteins [$A_{2958}/(A_{2852}+A_{2923})$], RNA/DNA ($A_{1121}/A_{1015}$), protein/carbohydrates (A_{1171}/A_{1152}), and RNA/carbohydrates [$(A_{1082}-A_{1056})/(A_{1028})$] ratios. This enabled further divisions of cancerogenous transformation of the cell lines into three premalignant stages according to the t-value of the discriminate classification score (DCS). Only the last stage of the disease progression exhibited morphological signs of malignancy, showing *de facto* an advantage of FTIR microscopy over standard techniques in the cell examination. In general, the analysis of biomarkers derived from spectra of cells and different neoplastic stages of cervical tissue (CINI/II/III) indicates the decrease in lipid, DNA, and carbohydrates content during the cancer progression, whereas the absorptions of proteins and RNA have a tendency to increase gradually. These results were found in the agreement with other studies on detection of malignancy from FTIR spectra. Moreover, this approach showed relatively high accuracy (*ca*. 90 %) and specificity (*ca*. 92 %) in classification of biopsies of cervical tissue from 37 patients diagnosed as invasive cervical carcinoma [104].

An attempt towards discrimination of a spectral phenotype HPV-infected cervical tissue has been undertaken by using ATR FTIR spectroscopy [105]. PCA-LDA analysis did not categorise low-grade from high-grade squamous intraepithelial lesions associated with HPV infection types. However, in a group of low-grade cytologically diagnosed with high-risk oncogenic HPV16/18, the multivariate analysis discriminated women based on their age into the two classes (20–29 years *vs*. 30–39 years). PCA loading plot exhibited the most prominent spectral differences between the both groups corresponding to vibrations of proteins (amide II and III bands, a band at 1404 cm^{-1}), phosphorylated proteins (972 cm^{-1}), and carbohydrates (1153 cm^{-1}). Spectral features specific for 4th decade women were those associated

with proteins (1404 cm^{-1}) and their phosphorylation [105]. More detailed modelling of the HPV effect on IR spectral profile of cell lines has been reported by Ostrowska et al. [106].

In a more recent study the effect of Mie-scattering on FTIR images cervical tissue was investigated [17]. Figure 15.8a shows uncorrected raw spectra selected from regions containing superficial exfoliating cells (yellow), intermediate epithelium (pink), basal layer (yellow) and stroma (blue). Figure 15.8b shows the same spectra after RMie-EMSC. The RMie-EMSC spectra are devoid of sinusoidal baseline modulation and the positions and intensities of the amide I mode are in the expected range for these cell types. Figure 15.8c shows the H&E section of the normal cervical epithelium; while Fig. 15.8d shows the corresponding amide I map showing protein concentration after RMie-scatter correction. Figure 15.8e shows a map derived from UHCA performed on the uncorrected spectra, while Fig. 15.8f shows the same map after RMie-EMSC has been applied. The UHCA map produced from the RMieS corrected FTIR image is similar to the uncorrected result but on closer inspection superior in that the epithelium shows improved layer demarcation into the exfoliating cell layer (gold), superficial layer of squamous epithelium (dark green), the intermediate layer (pink and some glycogen rich areas in blue) and the parabasal layer (orange). The stromal tissue is described by 3 clusters in the uncorrected UHCA map but only 2 clusters in the corrected UHCA map. The stroma is relatively homogeneous in terms of chemical content so the stroma being described by fewer clusters correlates more satisfactorily with the known properties of stromal tissue. These observations demonstrate that pre-processing FTIR image spectra with the RMieS correction prior to performing UHCA can improve the clustering correspondence with tissue features as well as achieve clustering parsimony.

Although the studies presented in this section are promising, the diagnosis of cervical cancer by the use of FTIR microspectroscopy is still a challenging task, since many factors can affect biochemical, and consequently spectral status of this tissue. Nevertheless, the potential application of FTIR spectroscopy to cervical cancer detection including cells and tissues is still an active research area [107–114]. Most of these studies have focused on correlating the spectral information with changes in biochemistry associated with the progression from healthy to malignant conditions and demonstrate the applicability of FTIR technique in tandem with proper statistical tools for future cancer diagnostic systems.

15.5.2 Brain Tissue

Since FTIR microspectroscopic analysis requires tissue sections, most of the human brain samples originate from biopsy (only tumours) or autopsy. This limits the applicability of FTIR imaging in diagnosis of diseases of the central nervous system. Despite this, the method has shown high sensitivity and specificity in grading brain tumours and its origin, even during surgery. Here, spectral changes are found in the entire mid-IR region of the spectrum revealing changes in intensity of

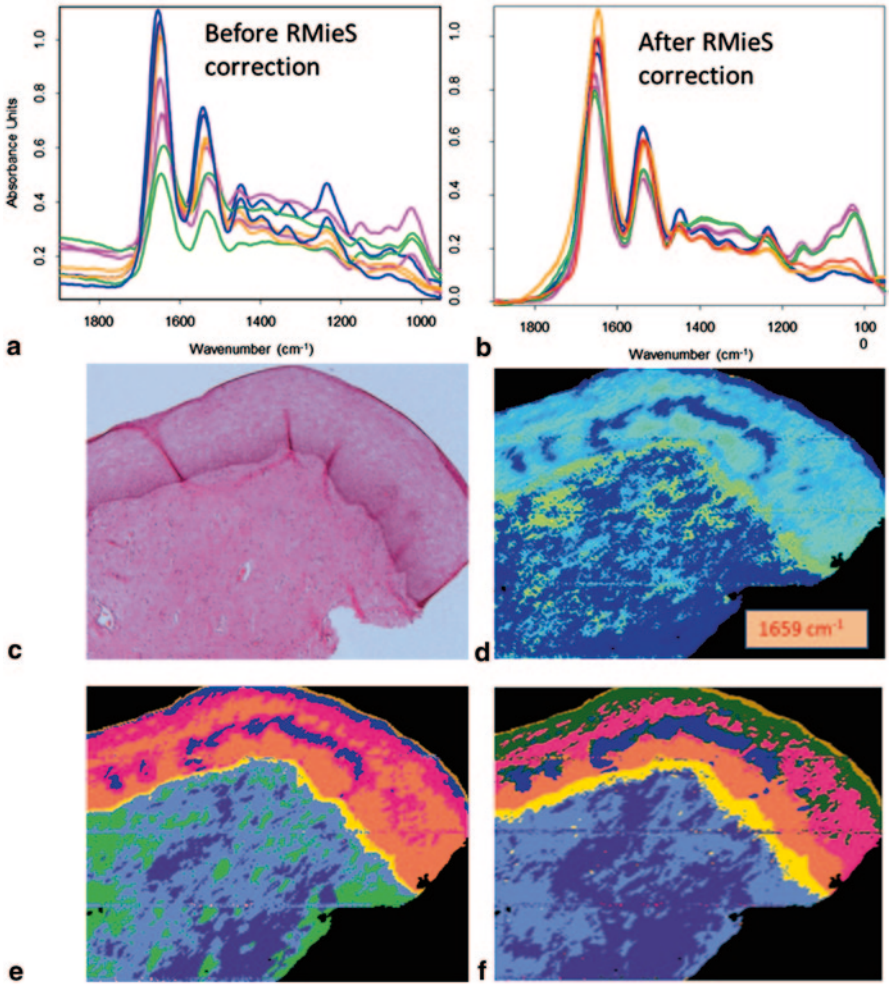


Fig. 15.8 **a** Spectra from different cell layers prior to RMie-EMSC. **b** After RMie-EMSC. **c** H&E stained section of the normal cervical epithelium. **d** Map generated by integrating the area beneath the amide I mode. **e** UHCA performed on data before RMie-EMSC. **f** UHCA performed on data after RMie-EMSC

carbohydrates moieties, lipid and protein content, cholesterol, collagen, phosphodiester-containing macromolecules and others [115–125]. The changes are associated with the type of the tumour and its origin. It should be noted that various multivariate approaches have been employed for sensitive and selective classification of the grade and origin of brain tumours, i.e. UHCA [120, 122], K-means [118, 121, 123], SIMCA (Soft Independent Modeling of Class Analogy) [119, 123], PCA [123, 125], LDA [116–118, 124] and their combinations.

In this section, we introduce the reader to the results of FTIR imaging of the brain tissue in neurodegenerative disorders such as Alzheimer's and Parkinson diseases, multiple sclerosis and others, for which the etiology or clinical diagnosis have not been clearly revealed, and are still the subject of intensive investigations. From this point of view, the clinical and therapeutic experiments on animal models and spectral characteristics of pathologically changed brain tissue of humans will be presented.

Reports on normal brain tissues have demonstrated the ability of FTIR spectroscopy to observe the two main brain tissue types, the grey and white matter, that are composed of unmyelinated neurons (neural cell bodies) and myelinated axons, respectively [115, 118]. The myelin sheath around the axon is a lipid-rich layer (mainly a glycolipid—galactocerebroside). FTIR spectra of the white matter exhibits high absorbance of the lipid vibrations, including bands at *ca*. 2922, 2852, 1740, 1468, and 1380 cm^{-1} attributed to vibrations of the CH_2 groups as well as *ca*. 1238 and 1085 cm^{-1} originating from the phosphodiester modes. Figure 15.9a depicts an H&E stained coronal section of normal rat brain tissue showing both the left and right hemisphere along with a UHCA performed using 8-clusters (15.9b). At eight clusters many anatomical features can be identified in the normal brain sample. Five layers are distinguishable in the cortex (yellow–orange–brown–green–brown) that may correspond to the outermost Brodmann layers. Three of the eight clusters (very-pale blue, red and yellow) corresponded to poor quality spectra that exhibited strong evidence of Mie scattering. The five remaining clusters correspond well with the tissue structures as follows; a cluster corresponding to the Brodmann layer IV in the cortex (green), a cluster containing an inner layer of the cortex and the grey matter areas in the caudoputamen (brown), a cluster for the white matter spots in the caudoputamen and most of the basal ganglia (orange), a cluster located in the lateral portions of the corpus callosum (light blue), and lastly, a cluster located at the medial portion of the corpus callosum, the commissure anterior and a medial spot in the basal ganglia (dark blue).

Figure 15.10 shows a H&E section of a section of rat brain tissue with fully developed tumour along with 2 cluster and 8 cluster UHCA maps. It is clear that in only 2 clusters the tumour is totally delineated from the rest of the brain. Moreover, one can observe the tumorous cells linking the primary to the secondary cluster, which is not resolved in the H&E section demonstrating the power of FTIR imaging combined with UHCA.

Neurodegenerative pathologies such as Alzheimer's and Parkinson's diseases, multiple sclerosis, prion diseases, and epileptic seizures seem to result from similar cellular and molecular mechanisms, since lipid peroxidation, the formation of inclusions, and protein aggregation are correlated with a change in secondary structure, which are often associated with disorders of the brain tissue.

One of the hypotheses explaining the cause of Alzheimer's disease (AD) is extracellular deposition of β-amyloid peptide (Aβ) in the grey matter of the brain. The formation of neuritic plaques are associated with folding of Aβ into a β-sheet conformation. In 1996, Choo and co-workers demonstrated the first attempt at detecting β-amyloid protein in human AD brain tissue by using synchrotron infrared

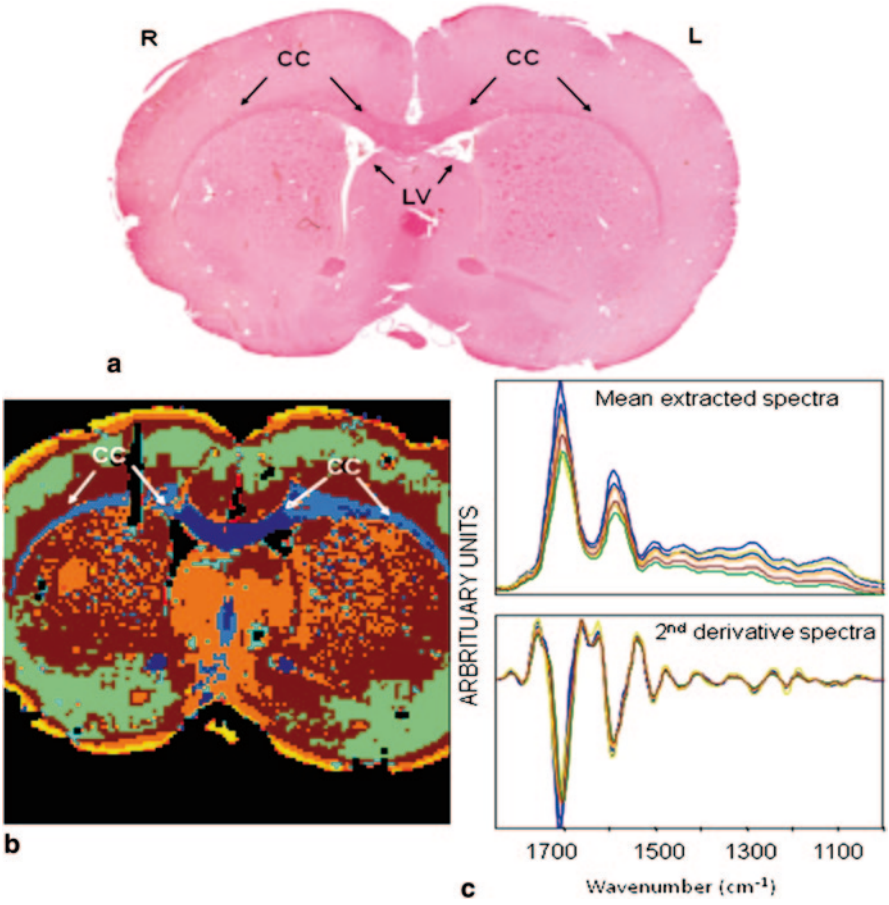


Fig. 15.9 **a** H&E coronal sections of healthy rat brain. **b** UHCA map generated for eight clusters using the 900–1800 cm^{-1} spectral window, in the healthy brain. **c** Mean extracted spectra with their 2nd derivatives. The colours are explained in the main text

microspectroscopy [126]. FTIR mapping of a tissue section derived from autopsy material showed a significant change in the amide I wavenumber in the region of neuritic plaques identified earlier from a histological examination. The position of this band shifts from 1652–1654 cm^{-1} in the periphery of the plaque to 1632–1634 cm^{-1} in its center. This clearly indicates a substantially different conformation of proteins in abnormal deposit in the AD tissue, suggesting an elevated level of β-sheet proteins. The position of the plaque protein found in the tissue differed from the position of synthetic amyloid (1620–1628 cm^{-1}), however, this fact can result from packing of other proteins around the amyloid deposit that may influence the structure of the aggregated β-amyloid [126]. The spectral features associated with an elevated content of β-sheet proteins in *ca*. 30 plaques of the human AD brain tissue has been confirmed by Miller et al. [127]. In this case, the position of the amide

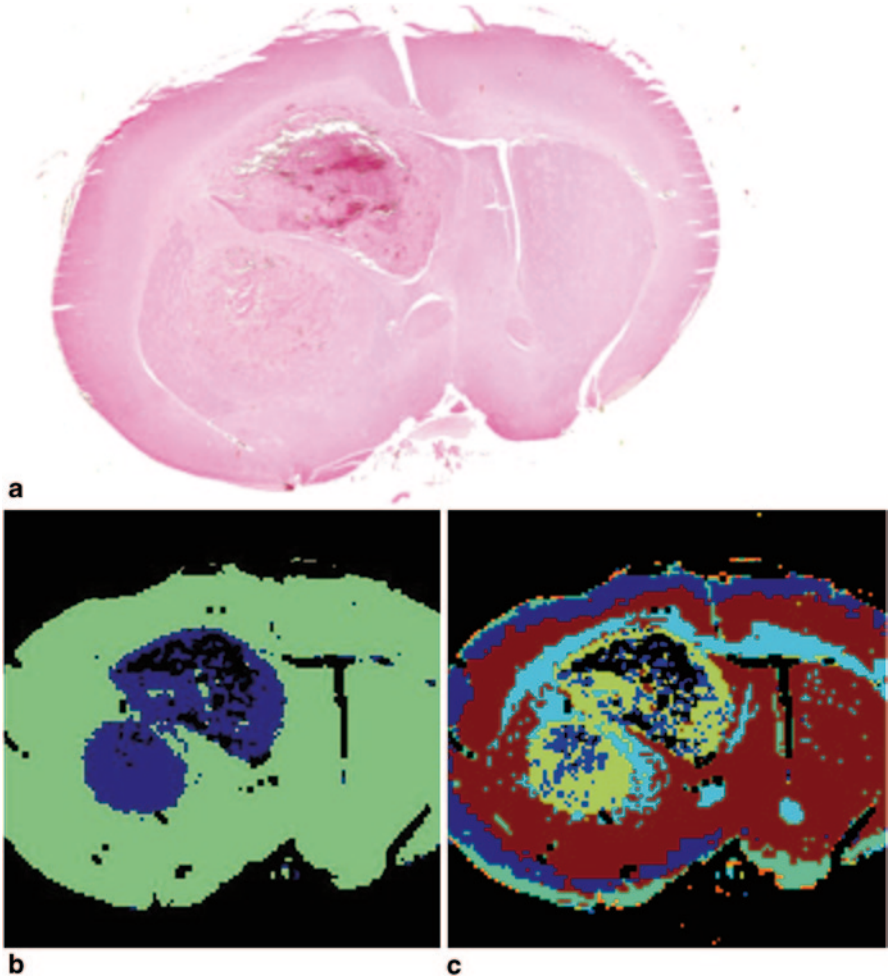


Fig. 15.10 **a** H&E section of rat brain showing a large tumour in the right hemisphere. **b** UHCA 2 cluster map clearly showing primary and secondary tumour. (C) UHCA 8 cluster map showing primary and secondary tumour along with other histological features

I band within the plagues is found at similar wavenumber as previously reported and is accompanied by a shoulder feature at 1655 cm^{-1}. However, the FTIR spectrum of a pure peptide $A\beta_{1-42}$, the major constituent of AD plaques, shows the presence of an amide I band at 1625 cm^{-1}, confirming a heterogeneous composition of proteins within the plaque. In another experiment, synchrotron-based IR maps have been correlated with synchrotron X-ray fluorescence (SXRF) imaging of metal elements [127]. This study showed a high accumulation of Cu, Zn, and β-sheet proteins in the amyloid plaques of the human AD tissue. However, the same correlation performed for the brain tissue sections of double transgenic PSAPP mouse, which represents the end-stage of AD, does not reveal the same effect of metal accumulation

within the amyloid plaques [128]. This animal model expresses a chimeric mouse/human amyloid precursor protein gene and a mutant human presenilin 1 gene, allowing the mice to secrete a human β-amyloid peptide. Although FTIR imaging confirms an increase in the protein content within the plaques by 61 % in comparison to the surrounding region, only a 29 % increase in Zn was observed compared to Cu (less by 22 %), Fe (less by 17 %), and Ca (less by 7 %). In the human AD tissue, the concentration of metal ions increased by 908, 1171, 573, and 9838 %, respectively. This suggests that metals may exhibit a weaker influence on the progression of AD in transgenic mouse than in human AD, and therefore this model may be a limited representation of human AD pathology [128]. Despite this, the PSAPP mouse model has been used in the FTIR examination of unsaturated lipid levels in the hippocampus tissue during aging (13–56 weeks) [129]. The presence of these macromolecules affects some functions of the brain such as membrane fluidity or signal transduction, while their degradation may signal lipid peroxidation. Firstly, UHCA of recorded IR images divided the hippocampus section into the three subregions of interest corresponding to the axonal, dendritic, and somatic layers. Average spectra extracted from each layer revealed the decrease in content of lipids in this order. Next, PCA analysis showed the disappearance of the unsaturated lipid band centred *ca*. 3015 cm^{-1} in the axonal layer for the PSAPP group compared to the control group for 56 week-old animals. Tracking the changes in intensity of this band reveals an increase in lipid unsaturation in the control mice with age, but not in the AD transgenic group. This fact indicates that the accumulation of unsaturated lipids in healthy animals is associated with the alteration of fatty acid composition during brain development. In turn, the deficiency in these macromolecules in the AD group may be explained by biochemical processes such as impaired biosynthesis of long-chain polyunsaturated fatty acids during the disease progression or Aβ-mediated lipid peroxidation. This study also shows that the AD pathogenesis can occur in the white matter of the brain too [129]. Other spectral markers of AD diseased tissue are proposed from an animal model experiment, in which the hippocampal, cortical, and caudal tissues of transgenic TgCRND8 mice were examined by synchrotron FTIR mapping and correlated with Raman spectroscopy [130]. This animal model expresses double mutant amyloid precursor protein and displays dense-cored plaques in the brain. The analysis of IR spectra showed the presence of a sharp band at 1304 cm^{-1} in the TgCRND8 tissues, which corresponded to localisation of crystalline inclusions. This band is assigned to creatine and its highest content is observed in the hippocampus of an advanced AD brain (the 72-week-old transgenic mice). The creatine band is also observed in IR images of *post mortem* human tissue of the hippocampus originating from two Alzheimer diseased patients [130]. The presence of creatine is associated with reduced metabolism of the AD brain, since this molecule with phosphocreatine regulates energy balance in the brain. This process is observed in the dementia-related disorders and may indicate a disturbed function of mitochondrial creatine kinase catalysing creatine phosphorylation [130]. A few creatine deposits can appear in normal ageing (in > 14-month mice), but more than 30 deposits randomly distributed in the hippocampus were already detected in the 5-month TgCRND8 brain, and this number increased

significantly with ageing of animals with AD disease [131]. The comprehensive analysis of IR images recorded for the same animal model has given an insight into chemical composition of the two types of the Aβ deposit, namely dense-core and diffuse amyloid plaques [132]. The former is associated with dystrophic neurites and activated glia cells, while the diffuse plaques are amorphous deposits of β-amyloid peptide without a core and dystrophic neurites. Dense-core plaques are mainly found in the cortex and hippocampus of the TgCRND8 mice and they are characterized by the increased intensity of the amide I band attributed to the presence of β-sheet proteins [132]. Moreover, IR imaging of the periphery of the dense-core plaques showed an elevated level of phospholipids. This increase was most pronounced around the large deposits, suggesting location of this plaque within a phospholipid envelope [132]. Interestingly, synchrotron/FPA-based IR imaging of the TgCRND8 hippocampus has demonstrated the ability of this instrumental system to image single neurons located in the close proximity of the amyloid plaque [133]. IR images recorded with a high spatial resolution show that phospholipids permeate the core of the plaque, but β-sheet aggregation of the amyloid peptide extends into the phospholipid bilayer. In turn, the diffuse plaques identified firstly by immunostaining do not show distinct spectral features when compared to the remaining area of the tissue section [132]. The amide I band appears at a wavenumber value typical for α-helical conformation of proteins and no other bands associated with other biomoleules change their positions or relative intensities. This fact is associated with unaltered morphology of the diffuse plaques, and consequently with unchanged chemical composition. This type of plaque appears in both normal and AD aged brain, indicating that separate processes are responsible for its formation, whereas the core-dense plaques are typical for the AD disorder [132]. Interestingly, an animal model of pilocarpine-evoked epilepsy shows spectral features similar to Alzheimer's disease [134, 135]. IR imaging of a sector 3 in the Ammon's horn (CA3) and dentate gyrus (DG) of the hippocampus revealed the presence of the amide I band at 1657 and 1631 cm^{-1} in the control and diseased tissues, respectively. This shift is accompanied by a significant decrease in the absorbance at 1657 cm^{-1} [134]. Protein conformation changes towards β-sheet structures appear along with an increase in content of saturated phospholipids in the pathological hippocampus, especially in pyramidal cells of CA3 as well as in multiform and molecular layers of the DG. The increased level of saturated lipids likely results from lipid peroxidation and it is proposed as a possible marker of oxidative stress, a common biochemical event in many neurodegenerative disorders. The appearance of protein and phospholipid structural changes together may also be indicative of a process responsible for seizure activity [135]. In addition, creatine inclusions accompanied the β-sheet protein deposits [134]. Few inclusions were observed in the healthy brain sections, whereas the number of creatine deposits increased to 100 in the area of the DG and the multiform hippocampal at the acute phase of status epilepticus. It is worth-mentioning that creatine deposits were also identified by the FTIR imaging technique in the human spinal cord, brain stem, and motor neuron cortex of two amyotrophic lateral sclerosis human cases, but no alternation of protein conformation was noted [136].

Preliminary FTIR studies on biochemical changes observed in the human brain in response to Parkinson's disease (PD) have focused on nerve cell bodies in the substantia nigra [137]. Degeneration and loss of dopamine-producing nerve cell bodies in PD mainly appear in this part of the brain. In general, the spectral region of 1380–1000 cm^{-1} (nucleic acids, carbohydrates, and amide III) shows the most pronounced differences between control and pathological neurons. Additionally, a variation in secondary structure of proteins is indicative of abnormal protein aggregation, which may be associated with a special role of the α-synuclein in the aggregation process, since this protein constitutes a major component of Levy bodies characteristic of Parkinson's disease. The examination of the amide I region exhibits the shift of the most intensive amide I band from 1656 to 1662 cm^{-1} in the control and diseased substantia nigra, respectively, along with the presence of the two shoulders at 1682 and 1643 cm^{-1} in the PD tissue. The authors associated these spectral changes with appearing β-turn and β-sheet proteins in PD disorder of neurons [137].

The FTIR profile of plaques found in the white matter of a chronic multiple sclerosis (MS) case in a human is associated only with the loss of phospholipids during demyelination of the myelin sheet [138]. The most pronounced spectral changes observed in the subcortical white matter of MS tissue sections derived from autopsy material reveals an increase in the 1740 cm^{-1} absorbance band, along with the reduction of the overall lipid content and broadening of the amide I band. The latter is slightly shifted from 1660 cm^{-1} in the control group to 1657 cm^{-1} in MS tissues. Such changes may be associated with lipid and protein oxidation induced by reactive oxygen and nitrogen intermediates. Thus, a role of free radicals in the etiology of MS has been implicated, since the oxidation products are found in MS lesion sites only. The interpretation of the observed spectral changes has been supported by *in vitro* FTIR studies of lipid and protein oxidation processes [139]. FTIR imaging has been successfully used in monitoring chemical changes during the progression of multiple sclerosis as well as after its prevention [140].The experiment was performed for an animal model, experimental autoimmune encephalomyelitis (EAE), typical in studies on the MS etiopathology and treatment. IR images of cerebellum tissues collected during the progression of the disease confirms earlier results that there is a dramatic decrease in lipids located within sites of the MS lesions. In addition, the images revealed that the formation of MS lesions is accompanied by increasing content of nucleic acid (based on the increased intensity of the 965 cm^{-1} band). These findings were in agreement with immunofluorescence staining for myelin basic protein. Discrimination of tissue layers and lesions required the use of PCA and ANNs that indicated lipid, protein, and nucleic acids bands are spectral markers for MS. The ANN approach enabled identification of early lesion formation before it could be clinically recognized. Interestingly, FTIR imaging indicates chemical changes in EAE-protected mice by vaccination with Nogo-A. In particular the cerebellums of fully-protected mice showed alterations in myelin protein structure. Namely, α-helical proteins are found to be typical for the vaccinated EAE animals, while a higher content of proteins with antiparallel β-pleated structures occurs in the diseased mice. The amide I bands are found at 1654, 1690, and 1635 cm^{-1},

respectively. This may indicate that a process of protein conformational change is concomitant with demyelination following Nogo-A vaccination [140].

The family of transmissible spongiform encephalopathies (TSE), called prion diseases, is a class of neurodegenerative diseases also associated with the presence of misfolded protein—prion protein PrP^{Sc} [141]. This belongs to the family of glycoproteins and appears in infected neurons, however deposits of PrP^{Sc} are not uniformly distributed throughout neurons, even in the terminated state of the disease. FTIR microspectroscopy studies have been focused on scrapie-infected hamsters only and they have evidenced protein changes in neurons by the observation of an IR band in the region of β-sheet conformation [142–144]. Worth mentioning is the fact that no relevant spectral differences in the amide I position were found in IR spectra recorded with the use of a globar source and single element detectors, but only utilizing synchrotron IR radiation revealed changes in protein conformation [144]. This was explained by the higher spatial resolution of synchrotron-based measurements that enabled detection of small-size PrP^{Sc} deposits. The synchrotron-based IR spectra show that the α-helical amide I band at 1657 cm^{-1} decreases in the infected hamsters along with a band appearing at 1637 cm^{-1}. An additional peak at 1631 cm^{-1} was also observed in some neurons. The examination of over 10,000 IR spectra of control and diseased neuronal cells indicated that the majority of the investigated cells show α-helical conformation with only a minority exhibiting a high content of β-sheet structure in the control group in contrast to the scrapie-infected animals. Since immunostaining with the PrP specific antibody 3F4 excludes the presence of misfolded proteins in the healthy tissues, the high concentration of β-sheet proteins found in the infected brain tissues is proposed as an indicatior of PrP^{Sc} deposition [144].

15.5.3 Cardiac Tissue

Infrared microspectroscopic imaging has been applied to examine the extracellular matrix of the myocardium in cardiomyopathy (CMP). Since diffuse and focal fibrosis is believed to accompany congestive heart failure, these studies have focused on searching for the most abundant collagen type I and III deposits in the ventricular tissue. Preliminary studies on spectral features of this tissue have demonstrated an elevated content in collagen [145, 146]. Characteristic IR maxima of the healthy rat heart tissue were observed in the region of carbohydrate residues attached to collagen (1031 and 1082 cm^{-1}), amide III bands of collagen (1204, 1240, and 1283 cm^{-1}), and the CH_2 wagging vibrations (1315 and 1338 cm^{-1}). In addition, the 3_1 helical structure of collagen of type I gives amide I bands at 1670 and 1635 cm^{-1}, that are assigned to weakly hydrogen-bonded proline and hydroxyproline residues and strongly hydrogen-bonded glycine carbonyls, respectively [145]. Thus, the normal extracellular matrix of the myocardium can be discriminated from the diseased tissue by spectral variations of the collagen spectral features. The first findings from FTIR mapping of the ventricles of control and cardiomyopathis hamsters have

revealed increasing intensity of the 1204 and 1240 cm^{-1} bands in diseased tissue [146]. Two-dimensional maps for the band at 1204 cm^{-1} correlated well with trichrome-stained collagen deposition. The data indicates heterogeneous distribution of collagen with its moderate increase in cardiomyopathic hamsters. The collagen depositions were mainly localised in the left ventricle [146].

Next, experimental cardiomyopathy induced by chronic AT_1 receptor blockade has been used to illustrate the spectroscopic differences among control, untreated cardiomyopathic and losartan-treated cardiomyopathic hearts, particularly in an area of micro-infarctions throughout both the ventricular tissues [147]. This animal model mimics the gradual progression of congestive heart failure in humans. Synchrotron-based FTIR instrumentation was used to improve the spatial resolution of the collected maps. An examination of the amide I and III regions shows the presence of the collagen bands listed above in most cardiomyopathic tissues and in some losartan-treated tissues on the contrary to normal cardiac tissue (rarely detected). In addition, the comparison of collagen and lipid distributions for the three models shows a reduction of lipid content in losartan-treated cardiomyopathic tissue in 200 day-old animals. This is concomitant with the replacement scar tissue of the infarction regions. Furthermore, crystalline deposits surrounded by collagen are detected in the necrotic regions. Based on their Raman spectra, crystals were readily identified as calcium hydroxyapatite responsible for mineralisation in the heart tissue. The collagen content was quantified by a number of pixel spectra, which showed evidence of collagen bands. In the early stage (65-day animals), about 2.5% of the total mapped area exhibited the characteristics of the necrotic tissue in the three groups, despite the fact that both the control and losartan-treated cardiomyopathic tissues were completely free of visible fibrosis. This was explained by a cardioprotective effect from the treatment. In the 200-day tissue sections, collagen-containing spectra were found in 2, 27 and 12% of the total number of spectra from control, diseased, and treated hearts, respectively, indicating reduction of fibrosis due to successful treatment in late-stage cardiomyopathic hamsters [147].

FTIR monitoring of the treatment of infarcted rat myocardium by targeted delivery of vascular endothelial growth factor (VEGF) has been reported by Cheheltani et al. [148]. Here, collagen deposition in the infarct zone of the left ventricles was mapped using the integrated intensity of a band at 1338 cm^{-1}, because only this IR maximum was found to be specific for collagen in the investigated tissue sections [148]. The maps were then compared with immunostaining for collagen I by Alexa Fluor 594 donkey anti-mouse IgG and the two histological staining protocols by using trichrome and picrosirius red, which also illustrated the distribution of collagen of type I. Correlation coefficient values of *ca*. 0.8–0.9 indicated a very good agreement between collagen distribution derived from the IR and staining images. This approach was also employed for the comparison of the infarcted myocardium with the VEGF-treated group. In this case, the size of collagen scarring decreased from 18.7 to 1.6% as revealed by the FTIR technique and was comparable to the staining results. These findings demonstrate that FTIR imaging can be a powerful label-free tool in the evaluation of collagen deposition and remodelling after myocardial

infarction [148]. The observation of collagen fibrosis by using FTIR imaging has been also reported for the cardiomyopathy induced by an infection with coxsackievirus, recognized as a highly active etiological agent in the development of myocarditis [149]. FTIR images constructed for a collagen band and lipid/protein ratio show similar results to those observed for cardiomyopathic hamsters [147]. No differences are observed for early, chronic, and acute stages of the infection [149]. All of them exhibited *ca*. 26% increase in the collagen content along with a significant reduction of lipids in comparison to the control group. This may suggest a similar etiology of cardiac dysfunction with a possible genetic and viral origin [149].

Other experiments have linked diabetes of type I [150] and early life growth restriction [151] with alternations of biochemical composition of the myocardia. Both the experiments have shown different spectroscopic variations across the entire spectral region in the disordered myocardia. This confirms that FTIR spectroscopy is a sensitive tool in detecting and/or monitoring various cardiac diseases and/or their pathogenesis. In the case of diabetic-induced changes in rat hearts, IR mapping was performed for both ventricles and the vein in the left ventricle. In the diseased myocardium, some spectral changes were found to be specific for the ventricles or veins only, whereas other spectral changes exhibit a similar pattern in all the investigated areas. For instance, an increase in content of olefinic acids (the 3012 cm^{-1} band) and collagen (the 1201 cm^{-1} band) were only observed in the vein of the left ventricle. Whilst the calculated lipid/protein ratio increases in the entire area of the myocardium in diabetic rats, but it is smaller in the right ventricle than in the left one and its veins. The authors suggested that the accumulation of lipids due to damage of the mitochondria in diabetic 5-week old rats results from improper metabolism of cardiac lipids. Such a lipid accumulation has been previously observed for diabetic 14-week old rats. Furthermore, FTIR spectra also revealed variation in protein secondary structure. The amide I band attributed to α-helical structures is present at *ca*. 1656 cm^{-1} in the healthy myocardium, whereas the second band at 1637 cm^{-1}, typical for β-pleated sheet conformations appears in the diabetic tissue. A significant increase of the latter is observed in the left ventricle, while reduction of the α-helix band intensity was found to be specific for the right ventricle. These conformational changes are presumed to result from the covalent attachment of glucose to amino groups of proteins that consequently lead to the rearrangement of protein structures. There is also an apparent glycogen accumulation accompanied by a reduction in content of cholesterol esters in the diabetic hearts, since the 1151/1171 cm^{-1} ratio decreases in the diabetic group with respect to the control one. In addition, the increase in intensity of a band at 1041 cm^{-1} indicates an accumulation of oligo- and polysaccharides [150].

Zohdi et al. have demonstrated the effect of a low protein diet during pregnancy and lactation on biochemical changes in the left ventricle of the rat myocardium [151]. FTIR false colour images were constructed by using UHCA in the fingerprint region, while the total protein and lipid content was assessed by integration of the amide II and C=O stretching bands, respectively. In addition, a band at 1228 cm^{-1}, called the "mixed" band, exhibited contributions from the amide III mode of proteins, the $\nu_{as}(PO_2^-)$ vibration of phospholipids and/or phosphorylated proteins, and

the $\nu_s(SO_4^{2-})$ mode of proteoglycans. Interestingly, the protein, nucleic acids and collagen content does not change in intrauterine growth-restricted rats in contrast to lipids where the concentration is significantly higher compared to the control group. On the other hand, the characteristic glycogen triplet randomly appears with a higher intensity for the low-protein diet, probably indicating variation in glycogen accumulation in those hearts. Moreover, the spatial distribution for lipids and the "mixed" contribution band follows the protein content in contrast to the carbohydrates. Because the intensities of collagen and nucleic acid bands do not alter within control and disordered hearts, the 1228 cm^{-1} band is proposed as a marker of proteoglycans. This class of glycosylated proteins is found in normal cardiac tissue and its deposition affects cardiac performance and vascular functions. Since their distribution is similar to the distribution of proteins and lipids, the proteoglycan accumulation might be associated with upregulation of lipids in maternal protein restriction. In turn, carbohydrate deposition can result from glucose-impaired metabolism during intrauterine growth restriction. In this process, the excess of glucose is stored in tissues in the form of glycogen/starch and lipids [151].

15.5.4 Liver Tissue

The first reports on spectral characteristics of the liver tissue in its normal, cirrhotic, and neoplastic states demonstrated a good correlation between infrared microspectroscopy and pathological/histochemical results and showed bands indicative of the diseased state [152, 153]. The liver lobule contains columns of hepatocytes, 30–40 μm in diameter, and epithelial cells, which are rich in glycogen. The distribution and concentration of glycogen vary within the lobule, and can increase in some abnormal states. This is clearly shown in IR spectra from different points of the cirrhotic liver tissue section [152] and thus the variation of intensity of the glycogen bands at 1025, 1080, and 1152 cm^{-1} cannot be associated with the diseased state of the liver. The normal human tissue from an autopsy material did not reveal the presence of glycogen. The presence/absence of glycogen in the investigated tissues is consistent with PAS (periodic acid-Schiff) staining, which is sensitive to cellular glycogen. In the case of healthy tissue, the absence of glycogen results from metabolic activity *post mortem* leading to complete remove of this biomolecule from the liver. However, a direct correlation between the lack of glycogen and pathological state must be performed with caution, since a change in its concentration can occur due to normal biochemical processes and not pathology. For instance, studies on liver tumours in rats have suggested that some preneoplastic hepatocytes store an excess of glycogen [152]. Glycogen-free IR spectra of normal, cirrhotic, and neoplastic liver tissues also exhibit small spectral differences between the type of tissue in the region of 1100–900 cm^{-1} [152, 153]. The comparison of healthy and neoplastic liver samples demonstrates that the neoplastic tissue is characterized by a broad IR maximum at 1080 cm^{-1} along with more intense bands at 960 and 1240 cm^{-1} [152, 153]. All the bands are attributed to DNA vibrations and their appearance

is consistent with observations found in the cervical tissue and other cancerous tissues, as discussed in Sections 1 and 5.1. Here, an increase of DNA content was confirmed by MGP (methyl green-pyronin) staining for DNA and RNA as well immunohistochemical staining for Ki-67 protein—the marker of proliferative activity in human cells [152]. The first staining protocol shows the predominant localization of RNA in the rough endoplasmic reticulum and nucleoli in the normal tissue, while DNA is found in nuclei. Staining with Ki-67 demonstrates that *ca*. 30% of cells are proliferating in the neoplastic tissue. Neoplastic cells typically contain more nuclear DNA/RNA as a result of genetic and chromosomal interactions [152]. The detailed IR study on liver tissues deglycogenized by α-amylase and digested by RNase and DNase confirmed these results [153]. Briefly, normal and cirrhotic liver tissues display predominantly RNA spectral features (the asymmetrical shape of the 1080 cm^{-1} band), whereas the symmetrical DNA band shape dominates in neoplastic tissue in the region below 1100 cm^{-1}.

FTIR imaging has been successfully employed in detecting spectral markers that characterise the progression of liver fibrinogenesis [154], liver steatosis [155], and acetaminophen-induced hepatotoxicity [156]. Liver fibrosis is a wound-healing process, which leads to the formation of a fibrillar extracellular matrix in the liver as a consequence of injuries, including infections, toxins, and metabolic disorders [154]. The formation of fibrillar extracellular matrix is associated with the degradation and remodelling of the extracellular matrix, similarly to myocardiopathy discussed in Sect. 5.3. Although, there are several tests monitoring liver fibrosis, its early stage cannot be reliably detected. In this experiment, early fibrosis alternations in a rat bile duct ligation (BDL) model were firstly characterized biochemically by elevated levels of ALT, AST, and bilirubin in serum [154]. These results suggest the dynamic deterioration of hepatocellular function. Routine H&E staining shows progressive deposition of extracellular matrix surrounding the newly formed bile ducts after 2 weeks of BDL. However, mild bile duct proliferation was evident after the first week. Immunofluorescence staining of collagen and IR spectra from the liver sections confirmed the H&E results. The collagen band at 1338 cm^{-1} illustrated fibrosis distributions throughout the sections. The IR false colour maps accurately matched the collagen distribution recorded in the immunofluorescence photomicrograph, even though the spatial resolution of both the methods differ significantly, *ca*. 10 and 0.5 μm, respectively [154].

The combination of the two imaging techniques, FTIR and ToF-SIMS, has been employed to understand mechanisms underlying fatty liver diseases and to identify markers for diagnosis and prognosis of steatosis [155]. This study has been focused on detection of variations in lipid composition and metabolic abnormalities in steatotic and non-steatotic hepatocytes. The tissue sections were derived from the non-tumoral part of liver resection specimens. The comparison of the non-steatotic and steatotic vesicles clearly shows the absence of amide I and II bands in the steatotic regions. The IR spectrum of the fatty liver area is typical for lipids and reveals the presence of the unsaturated fatty acids only in this region of the liver. The IR false colour maps shows that the steatosis vesicles are surrounded by the protein-rich tissue and their lipid conformation dramatically differs from structure

and composition of lipids in the tissue. The ToF-SIMS analysis in turn identified precisely some of the lipids present in the liver. Cholesterol, monoacylglycerols and diacylglycerols (DAG) were detected as highly abundant species. Whilst saturated DAG species were detected mostly outside of steatotic vesicles, a high content of cholesterol along with DAG containing at least one unsaturated acyl chain is selectively concentrated in the vesicles, confirming the FTIR results. Despite the fact that the H&E histological examination of the normal and steatotic liver showed similar morphological features of non-steatotic regions in both samples, the FTIR spectra revealed important biochemical changes. Non-steatotic hepatocytes of the fatty liver are characterized by the higher intensity of the stretching modes of lipid and carbohydrate moieties. The ToF-SIMS imaging of the same region of the samples demonstrated the presence of small lipid droplets (< 10 μm) containing DAGs. These droplets are likely to correspond to the first event of lipid accretion. Thus, the combination of the two imaging techniques demonstrates early metabolic disturbance in fatty acids of the liver. In conclusion, infrared microspectroscopy hinted at global chemical changes that could be further investigated by the selective and sensitive mass spectrometry [155].

A detailed study combining FTIR imaging, histological analysis and sera markers of the liver function, on mouse liver injury induced by a popular pain-relief APAP has been undertaken by Gautam et al. [156]. The FTIR spectral changes were detected in the liver tissue after 0.5 h of oral drug administration. They revealed a significant lowering of the glycogen content accompanied by an increase in the concentration of cholesteryl esters and in the DNA/RNA ratio. This was highlighted by the intensity variation of bands at 1171 and 966/996 cm^{-1}, respectively. The changes of glycogen and cholesteryl esters were sustained up to 6 h post APAP treatment, while the DNA/RNA ratio increased significantly after 0.5 h, peaking at 1.5 h, thereafter was sustained up to 6 h. These changes correlated with a drop in total GSH and an increase in malonaldehyde, but the increase in ALT amounts and histological evidence of the liver injury occurred only after 3 h period and peaked with maximum liver damage by 6 h. Interestingly, the spectral changes found in the tissue were similar to spectral features of sera, indicating a possible application of FTIR spectroscopy in monitoring of hepatoxicity. Additionally, FTIR data from imaging of the tissue showed molecular changes due to pre- and post-treatment of the APAP treated mice with an antidote—L-methionine. ALT amounts indicated inhibiting the hepatotoxic effect, but FTIR spectra confirmed that only in case of L-methionine pre-treatment. The glycogen and DNA levels were restored to the level observed in the control group, while the content of cholesteryl esters remained elevated. This implicates that glycogen and DNA amounts might be indicative for the early rise of oxidative stress [156].

15.6 Outlook

The scene is now set for FTIR to play a significant role in pathological diagnosis; however, a number of hurdles need to be overcome prior to implementation into the clinical laboratory. One of the most important barriers to overcome is standardization of methodology, instrumental procedures and data processing. At present every lab working on FTIR imaging of tissue sections have their own method of sample preparation, instrumental parameters and spectral processing. These methods need to be standardized so a measurement recorded in Australia on one section would yield a similar result to the same section prepared and measured in Poland, for example.

The inability to detect strong contributions from A-DNA in fixed tissue sections is problematic given that the importance of DNA in the etiology of the disease. In the future it may be possible to fix B-DNA and resolve the important phosphodiester vibrations and other DNA marker bands. At present the only way to preserve B-DNA is to keep the sample in a hydrated state, which is clearly not conducive for routine diagnosis.

The contribution of light dispersion and Mie scattering to the spectral profile needs more investigation particular on how this effect influences the histological diagnosis. The question at present is do we correct for this light dispersion by performing RMie-EMSC (or some other approach) or do we leave the Mie scattering component in the spectrum because it too is diagnostically useful. If we consider that the size of the nucleus is one of the most important pathological criteria that a histologist uses to grade cancer then it seems counterintuitive that one would want to remove this information from the spectrum. The same argument can be made for dispersion. The optically dense regions of tissue are often the cancerous areas containing abundant nuclei. It is clear from the early studies applying FTIR imaging to diagnose CIN in cervical tissues that when the amide I and II region was used in the cluster analysis the best demarcation of cancer versus non-cancer was achieved [31]. This region of course is the most affected by dispersion as this is the region of highest absorbance in the FTIR spectrum. As biospectroscopists we try to achieve a “text book” spectrum devoid of baseline undulations and other physical contributions in pursuit of the pure chemical information, however, if our purpose is to develop diagnostic tools then one could argue that all contributions whether physical or chemical should be included in the ensuing analysis. The focus of future studies should address this central and pivotal question.

The other important consideration is choice of substrate given the inherent standing-wave-effect generated by IR trans-reflection slides. This effect can be reduced, in the case of tissue, if each section is evenly cut and a second derivative calculated

prior to image processing because the second derivative highlights band shifts and inflection points in the spectrum and does not rely on linear absorbance directly. In the case of fixed exfoliated cells, which vary in both size and thickness, the wavelength dependent scattering is exacerbated and more difficult to correct. It appears that IR transparent windows are the only solution to negate the standing-wave-effect in the case of fixed exfoliated or cultured cells although more work is required to verify the findings in relation to diagnostic capability.

FTIR imaging has come a long way in the last 10 years especially in terms of the data processing speeds and understanding contributions to the spectra. With the development of bench top synchrotron sources and refinements to the optic and detector electronics a new laboratory gold standard is on the horizon.

Acknowledgements Dr. B.R. Wood is financially supported by an Australian Research Council Future Fellowship FT120100926.

References

1. Mantsch HH, Jackson M (1995) Molecular spectroscopy in biodiagnostics (from Hippocrates to Herschel and beyond). J Mol Struct 347:187–206
2. Husseinzadeh N (2011) Status of tumor markers in epithelial ovarian cancer, has there been any progress? A review. Gynecol Oncol 120:152–157
3. Breen M (2010) Update on genomics in veterinary oncology. Top Companion Anim Med 24:113–121
4. Gould Rothberg BE, Bracken MB, Rimm DL (2009) Tissue biomarkers for prognosis in cutaneous melanoma: a systematic review and meta-analysis. J Natl Cancer Inst 101:452–474
5. Staudt LM, Dave S (2005) The biology of human lymphoid malignancies revealed by gene expression profiling. Adv Immunol 87:163–208
6. Moore JH, Asselbergs FW, Williams SM (2010) Bioinformatics challenges for genome-wide association studies. Bioinformatics 24:445–455
7. Tobin MJ, Puskar L, Barber RL et al (2010) FTIR spectroscopy of single live cells in aqueous media by synchrotron IR microscopy using microfabricated sample holders. Vib Spectrosc 53:34–38
8. Ellis DI, Dunn WB, Griffin JL et al (2007) Metabolic fingerprinting as a diagnostic tool. Pharmacogenomics 8:1243–1266
9. Lewis EN, Treado PJ, Reeder RC et al (1995) Fourier transform spectroscopic imaging using an infrared focal-plane array detector. Anal Chem 67:3377–3381
10. Bassan P, Lee J, Sachdeva A et al (2013) The inherent problem of transflection-mode infrared spectroscopic microscopy and the ramifications for biomedical single point and imaging applications. Analyst 138:144–157
11. Whelan DR, Bambery KR, Heraud P et al (2011) Monitoring the reversible B to A-like transition of DNA in eukaryotic cells using Fourier transform infrared spectroscopy. Nucleic Acids Res 39:5439–5448
12. Whelan DR, Bambery KR, Puskar L et al (2012) Quantification of DNA in simple eukaryotic cells using Fourier transform infrared spectroscopy. J Biophotonics doi:10.1002/jbio.201200112
13. Liu KZ, Tsang KS, Li CK et al (2003) Infrared spectroscopic identification of β-thalassemia. Clin Chem 49:1125–1132

14. Ami D, Neri T, Natalello A et al (2008) Embryonic stem cell differentiation studied by FTIR spectroscopy. Biochim Biophys Acta 1783:98–106
15. Choo LP, Wetzel DL, Halliday WC et al (1996) In situ characterization of β-amyloid in Azheimer's diseased tissue by synchrotron Fourier transform infrared microspectroscopy. Biophys J 71:1672–1679
16. Palombo F, Shen H, Benguigui LES et al (2009) Micro ATR-FTIR spectroscopic imaging of atherosclerosis: an investigation of the contribution of inducible nitric oxide synthase to lesion composition in ApoE-null mice. Analyst 134:1107–1118
17. Bambery KR, Wood BR, McNaughton D (2012) Resonant Mie scattering (Rimes) correction applied to FTIR images of biological tissue samples. Analyst 137:126–132
18. Noguchi T, Inoue Y, Tang XS (1997) Structural coupling between the oxygen-evolving Mn cluster and a tyrosine residue in photosystem II as revealed by Fourier transform infrared spectroscopy. Biochemistry 36:14705–14711
19. Maziak DE, Do MT, Shamji FM et al (2007) Fourier-transform infrared spectroscopic study of characteristic molecular structure in cancer cells of esophagus: an exploratory study. Cancer Detect Prev 31:244–253
20. Bogomolny E, Argov S, Mordechai S, Huleihel M (2008) Monitoring of viral cancer progression using FTIR microscopy: a comparative study of intact cells and tissues, Biochim Biophys Acta 1780:1038–1046
21. Ooi GJ, Fox J, Siu K et al (2008) Fourier transform infrared imaging and small angle x-ray scattering as a combined biomolecular approach to diagnosis of breast cancer. Med Phys 35:2151–2161
22. Chen L, Holman HYN, Hao Z et al (2012) Synchrotron infrared measurements of protein phosphorylation in living single PC12 cells during neuronal differentiation. Anal Chem 84:4118–4125
23. Benedeti E, Bramanti E, Papineschi F et al (1997) Determination of the relative amount of nucleic acids and proteins in leukemic and normal lymphocytes by means of Fourier transform infrared microspectroscopy. Appl Spectrosc 51:792–797
24. Salman A, Ramesh J, Erukhimovitch V et al (2003) FTIR microspectroscopy of malignant fibroblasts transformed by mouse sarcoma virus. J Biochem Biophys Methods 55:141–153
25. Mohlenhoff B, Romeo MJ, Wood BR et al (2005) Mie-type scattering and non-Beer-Lambert absorption behavior of human cells in infrared microspectroscopy. Biophys J 88:3635–3640
26. Wood BR, Chernenko T, Matthäus C et al (2008) Shedding new light on the molecular architecture of oocytes using a combination of synchrotron Fourier transform-infrared and Raman spectroscopic imaging. Anal Chem 80:9065–9072
27. Lasch P, Boese M, Pacifico A et al (2002) FT–IR spectroscopic investigations of single cells on the subcellular level. Vib Spectrosc 28:147–157
28. Ozek NS, Tuna S, Erson-Bensan AE et al (2010) Characterization of microRNA-125b expression in MCF7 breast cancer cells by ATR-FTIR spectroscopy. Analyst 135:3094–3102
29. Wrobel TP, Mateuszuk L, Chlopicki S et al (2011) Imaging of lipids in atherosclerotic lesion in aorta from ApoE/LDLR-/- mice by FT-IR spectroscopy and Hierarchical Cluster Analysis. Analyst 136:5247–5255
30. Chiriboga L, Xie P, Lee H et al (1998) Infrared spectroscopy of human tissue. I. Differentiation and maturation of epithelium cells in the human cervix. Biospectrosc 4:47–53
31. Wood BR, Chiriboga L, Yee H et al (2004) Fourier transform infrared (FTIR) spectral mapping of the cervical transformation zone and dysplastic squamous epithelium. Gynecol Oncol 93:59–68
32. Wood BR, Quinn MA, Tait B et al (1998) FTIR microspectroscopic study of cell types and potential confounding variable in screening for cervical malignancies. Biospectroscopy 4:75–91
33. Chiriboga L, Xie P, Zhang W et al (1997) Infrared spectroscopy of human tissue. III. Spectral differences between squamous and columnar tissue and cells from the human cervix. Biospectroscopy 3:253–257

34. Wood BR, Tait B, McNaugthon D (2000) Fourier Transform Infrared Spectroscopy as a Method for Monitoring the Molecular Dynamics of Lymphocyte Activation. Appl Spectrosc 54:353–359
35. Andrus PGL, Strickland RD (1998) Cancer grading by Fourier transform infrared spectroscopy. Biospectroscopy 4:37–46
36. Mordechai S, Sahu RK, Hammody Z et al (2004) Possible common biomarkers from FTIR microspectroscopy of cervical cancer and melanoma. J Microsc 215:86–91
37. Taillandier E, Liquier J (1992) Infrared spectroscopy of DNA. Methods Enzymol 211:307–335
38. Mohlenhoff B, Romeo M, Diem M et al (2005) Mie-type scattering and non-Beer-Lambert absorption behaviour of human cells in infrared micro-spectroscopy. Biophys J 88:3635–3640
39. Shimanouchi T, Tsuboi M, Kyogoku Y (1964) The structure and properties of biomolecules and biological systems. In: Duchesne J (ed) Advances in chemical physics, vol VII. Interscience, New York
40. Carr GL (2001) Resolution limits for infrared microspectroscopy explored with synchrotron radiation. Rev Sci Instrum 72:1613–1619
41. Sommer AJ, Katon JE (1991) Diffraction-induced stray light in infrared microspectroscopy and its effect on spatial resolution. Appl Spectrosc 45:1633–1640
42. Lasch P, Naumann D (2006) Spatial resolution in infrared microspectroscopic imaging of tissues. Biochim Biophys Acta 1758:814–929
43. Kazarian SG, Chan KLA (2010) Micro- and macro-attenuated total reflection Fourier transform infrared spectroscopic imaging. Appl Spectrosc 64:135A–152A
44. Filik J, Frogley MD, Pijanka JK et al (2012) Electric field standing wave artifacts in FTIR micro-spectroscopy of biological materials. Analyst 137:853–861
45. Holton SE, Walsh MJ, Bhargava R (2011) Subcellular localization of early biochemical transformations in cancer-activated fibroblasts using infrared spectroscopic imaging. Analyst 136:2953–2958
46. Amrania H, McCrow AP, Matthews MR et al (2011) Ultrafast infrared chemical imaging of live cells. Chem Sci 2:107–111
47. Kuimova MK, Chan KLA, Kazarian SG (2009) Chemical imaging of live cancer cells in the natural aqueous environment. Appl Spectrosc 63:164–171
48. Colley CS, Kazarian SG, Weinberg PD et al (2004) Spectroscopic imaging of arteries and atherosclerotic plaques. Biopolymers 74:328–335
49. Rogalski A (2005) HgCdTe infrared detector material: history, status and outlook. Rep Prog Phys 68:2267–2336
50. Norton P (2002) HgCdTe infrared detectors. Opto-Electron Rev 10:159–174
51. Miller LM, Smith RJ (2005) Synchrotrons versus globars, point-detectors versus focal plane arrays: selecting the best source and detector for specific infrared microspectroscopy and imaging applications. Vib Spectrosc 38:237–240
52. Diem M, Romeo M, Matthäus C et al (2004) Comparison of Fourier transform infrared (FTIR) spectra of individual cells acquired using synchrotron and conventional sources. Infrared Phys Technol 45:331–338
53. Bhargava R, Wall BG, Koenig JL (2000) Comparison of the FTIR mapping and imaging techniques applied to polymeric systems. Appl Spectrosc 54:470–479
54. Reffner JA, Martoglio PA, Williams GP (1995) Fourier transform infrared microscopical analysis with synchrotron radiation: the microscope optics and system performance (invited). Rev Sci Instrum 66:1298–1302
55. Miller LM, Dumas P (2006) Chemical imaging of biological tissue with synchrotron infrared light. Biochim Biophys Acta 1758:846–857
56. Tobin MJ, Chesters MA, Chalmers JM et al (2004) Infrared microscopy of epithelial cancer cells in whole tissues and in tissue culture, using synchrotron radiation. Faraday Discuss 126:27–39

57. Nasse MJ, Walsh MJ, Mattson EC et al (2011) High-resolution Fourier transform infrared chemical imaging with multiple synchrotron beams. Nat Methods 8:413–416
58. Dumas P, Jasmin N, Teillaud JL et al (2004) Imaging capabilities of synchrotron infrared microspectroscopy. Faraday Discuss 126:289–302
59. Carr GL, Chubar O, Dumas P (2005) Multichanel detection with a synchrotron light source: design and potential. In: Bhargava R, Levin I (eds) Spectrochemical analysis using multichanel detectors. Analytical chemistry series, chap. 3. Blackwell, Oxford, pp 56–84
60. Moss D, Gasharova B, Mathis YL (2006) Practical tests of a focal plane array detector microscope at the ANKA-IR beamline. Infrared Phys Technol 49:53–56
61. Petibois C, Cestelli-Guidi M, Piccinini M et al (2010) Synchrotron radiation FTIR imaging in minutes: a first step towards real-time cell imaging. Anal Bioanal Chem 397:2123–2129
62. Hirschmugl CJ, Gough KM (2012) Fourier transform infrared spectrochemical imaging: review of design and applications with a focal plane array and multiple beam synchrotron radiation source. Appl Spectrosc 66:475–491
63. Afseth NK, Kohler A (2012) Extended multiplicative signal correction in vibrational spectroscopy, a tutorial. Chemometr Intel Lab Systems 117:92–99
64. Lan T, Fang Y, Xiong W et al (2007) Automatic baseline correction of infrared spectra. Chin Opt Lett 5:613–616
65. Severcan F, Harris PI (2012) Vibrational spectroscopy in diagnosis and screening. In: advances in biomedical spectrosocpy, Vol. 6. IOS Press
66. Butler WL, Hopkins DW (1970) An analysis of fourth derivative spectra. Photochem Photobiol 12:451–456
67. Susi H, Byler DM (1986) Resolution-enhanced Fourier transform infrared spectroscopy of enzymes. Methods Enzymol 130:290–311
68. Savitsky A, Golay MJE (1964) Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Anal Chem 36:1627–1639
69. van de Hulst HC (1981) Light scattering by small particles. Dover, New York
70. Martens H, Naes T (1989) Multivariate Calibration. Wiley, Chichester
71. Martens H, Nielsen JP, Engelsen SB (2003) Light scattering and light absorbance separated by extended multiplicative signal correction. Application to near-infrared transmission analysis of powder mixtures. Anal Chem 75:394–404
72. Chen ZP, Morris J, Martin E (2006) Extracting chemical information from spectral data with multiplicative light scattering effects by optical path-length estimation and correction. Anal Chem 78:7674–7681
73. Kohler A, Sulé-Suso J, Sockalingum GD et al (2008) Estimating and correcting Mie scattering in synchrotron-based micorscopic Fourier transform infrared spectra by extended multiplicative signal correction. Appl Spectrosc 62:259–266
74. Martens H, Stark E (1991) Extended multiplicative signal correction and spectral interference subtraction: new preprocessing methods for near infrared spectroscopy. J Pharm Biomed Anal 9:625–635
75. Kohler A, Kirschner C, Oust A et al (2005) Extended multiplicative signal correction as a tool for separation and characterization of physical and chemical information in Fourier transform infrared microscopy images of cryo-sections of beef loin. Appl Spectrosc 59:707–716
76. Bassan P, Byrne HJ, Bonnier F et al (2009) Resonant Mie scattering in infrared spectroscopy of biological materials-understanding the 'dispersion artefact'. Analyst 134:1586–1593
77. Bassan P, Kohler A, Martens H et al (2010) Resonant Mie scattering (Rimes) correction of infrared spectra from highly scattering biological samples. Analyst 135:268–277
78. Bassan P, Kohler A, Martens H et al (2010) Rimes-EMSC correction for infrared spectra of biological cells: extension using full Mie theory and GPU computing. J Biophoton 3:609–620
79. Bassan P, Byrne HJ, Lee J et al (2009) Reflection contributions to the dispersion artefact in FTIR spectra of single biological cells. Analyst 134:1171–1175

80. Bassan P, Gardner P (2011) Biomedical applications of synchrotron infrared microspectroscopy: a practical approach, Moss D (ed). Royal Society of Chemistry, Cambridge, pp. 260–276
81. Bassan P, Sachdeva A, Kohler A et al (2012) FTIR microscopy of biological cells and tissue: data analysis using resonant Mie scattering (Rimes) EMSC algorithm. Analyst 137:1370–1377
82. Miljkovic M, Bird B, Diem M (2012) Line shape distortion effects in infrared spectroscopy. Analyst 137:3954–3964
83. Lasch P, Naumann D (1998) FT-IR microspectroscopic imaging of human carcinoma thin sections based on pattern recognition techniques. Cell Mol Biol 44:189–202
84. Lasch P, Haensch W, Lewis EN et al (2002) Characteristion of colorectal adenocarcinoma sections by spatially resolved FT-IR microspectroscopy. App Spectrosc 56:1–9
85. Fabian H, Lasch P, Boese M et al (2003) Infrared microspectroscopic imaging of benign breast tumour tissue sections. J Mol Struct 661–662:411–417
86. Ward JH Jr (1963) Hierarchical grouping to optimize an objective function. J Am Stat Assoc 58:236–244
87. Vapnik VN, Chervonenkis AYA (1971) On the uniform convergence of relative frequencies of events to their probabilities. Theor Probab Appl 16:264–281
88. Hush DR, Horne BG (1993) Progress in supervised neural network. IEEE Signal Proc Mag 10:8–39
89. Walboomers JM, Jacobs MV, Manos MM et al (1999) Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J Pathol 189:12–19
90. Poljak M, Kocjan BJ (2010) Commercially available assays for multiplex detection of alpha human papillomaviruses. Expert Rev Anti-infe 8:1139–1162
91. Wong PTT, Wong RK, Caputo TA et al (1991) Infrared Spectroscopy of Exfoliated human cervical cells: evidence of extensive structural changes during carcinogenesis. Proc Natl Acad Sci U S A 88:10988–10992
92. Wood BR, Quinn MA, Burden FR et al (1996) An investigation into FTIR spectroscopy as a biodiagnostic tool for cervical cancer. Biospectroscopy 2:143–153
93. Mordechai S, Sahu RK, Hammody Z et al (2004) Possible common biomarkers from FTIR microspectroscopy of cervical cancer and melanoma. J Microsc 215:86–91
94. Chiriboga L, Xie P, Lee H et al (1998) Infrared spectroscopy of human tissue. I. Differentiation and maturation of epithelium cells in the human cervix. Biospectroscopy 4:47–53
95. Chiriboga L, Xie P, Lee H et al (1998) Infrared spectroscopy of human cells and tissues: IV. Detection of dysplastic and neoplastic changes in human cervical tissue via infrared microscopy. Cell Mol Biol 44:219–229
96. Chiriboga L, Xie P, Vigorita P et al (1998) Infrared spectroscopy of human tissue: II. A comparative study of spectra of biopsies of cervical squamous epithelium and of exfoliated cervical cells. Biospectroscopy 4:55–59
97. Wood BR, Quinn MA, Tait B et al (1998) FTIR microspectroscopic study of cell types and potential confounding variable in screening for cervical malignancies. Biospectroscopy 4:75–91
98. Diem M, Chiriboga L, Lasch P et al (2002) IR spectra and IR spectral maps of individual normal and cancerous cells. Biopolymers 67:349–353
99. Chang JI, Huang YB, Wu PC et al (2003) Characterization of human cervical precancerous tissue through the Fourier transform infrared microscopy with mapping method. Gynecol Oncol 91:577–583
100. Wood BR, Chiriboga L, Yee H et al (2004) Fourier transform infrared (FTIR) spectral mapping of the cervical transformation zone and dysplastic squamous epithelium. Gynecol Oncol 93:59–68
101. Steller W, Einenkel J, Horn LC et al (2006) Delimitation of squamous cell cervical carcinoma using infrared microspectroscopic imaging. Anal Bioanal Chem 384:145–154

102. Mark S, Sahu RK, Kantarovich K et al (2004) Fourier transform infrared microspectroscopy as a quantitative diagnostic tool for assignment of premalignancy grading in cervical neoplasia. J Biomed Opt 9:558–567
103. Podshyvalov A, Sahu RK, Mark S et al (2005) Distinction of cervical cancer biopsies by use of infrared microspectroscopy and probabilistic neural networks. Appl Opt 44:3725–3734
104. Bogomolny E, Argov S, Mordechai S et al (2008) Monitoring of viral cancer progression using FTIR microscopy: a comparative study of intact cells and tissues. Biochim Biophys Acta 1780:1038–1046
105. Kelly JG, Cheung KT, Martin C et al (2010) A spectral phenotype of oncogenic human papillomavirus-infected exfoliative cervical cytology distinguishes women based on age. Clin Chim Acta 411:1027–1033
106. Ostrowska KM, Garcia A, Meade AD et al (2011) Correlation of p16INK4A expression and HPV copy number with cellular FTIR spectroscopic signatures of cervical cancer cells. Analyst 136:1365–1373
107. Cohenford MA, Godwin TA, Cahn F et al (1997) Infrared spectroscopy of normal and abnormal cervical smears: evaluation by principal component analysis. Gynecol Oncol 66:59–65
108. El-Tawil SG, Adnan R, Muhamed ZN, Othman NH (2008) Comparative study between Pap smear cytology and FTIR spectroscopy: a new tool for screening for cervical cancer. Pathology 40:600–603
109. Petter CH, Heigl N, Rainer M et al (2009) Development and application of Fourier-transform infrared chemical imaging of tumor in human tissue. Curr Med Chem 16:318–326
110. Rigas B, LaGuardia K, Qiao L et al (2000) Infrared spectroscopic study of cervical smears in patients with HIV: implications for cervical carcinogenesis. J Lab Clin Med 135:26–31
111. Sindhuphak R, Issaravanich S, Udomprasertgul V, Srisookho P, Warakamin S, Sindhuphak S, Boonbundarlchai R, Dusitsin N (2003) A new approach for the detection of cervical cancer in Thai women. Gynecol Oncol 90:10–14
112. Walsh MJ, Singh MN, Pollock HM et al (2007) ATR microspectroscopy with multivariate analysis segregates grades of exfoliative cervical cytology. Biochem Biophys Res Commun 352:213–219
113. Schubert JM, Bird B, Papamarkakis K et al (2010) Spectral cytopathology of cervical samples: detecting cellular abnormalities in cytologically normal cells. Lab Invest 90:1068–1077
114. Krishna CM, Sockalingum GD, Vadhiraja BM et al (2007) Vibrational spectroscopy studies on formalin-fixed cervix tissues. Biopolymers 85:214–221
115. Steiner G, Shaw A, Choo-Smith LP et al (2003) distinguishing and grading human gliomas by IR spectroscopy. Biopolymers 72:464–471
116. Beleteis C, Steiner G, Sowa MG et al (2005) Classification of human gliomas by infrared imaging spectroscopy and chemometric image processing. Vib Spectrosc 38:143–149
117. Krafft C, Thümmler K, Sobottka SB et al (2006) Classification of malignant gliomas by infrared spectroscopy and linear discriminant analysis. Biopolymers 82:301–305
118. Amharref N, Beljebbar A, Dukic S et al (2006) Brain tissue characterization by infrared imaging in a rat glioma model. Biochim Biophys Acta 1758:892–899
119. Krafft C, Shapoval L, Sobottka SB et al (2006) Identification of primary tumors of brain metastases by SIMCA classification of IR spectroscopic images. Biochim Biophys Acta 1758:883–891
120. Bambery KR, Schültke E, Wood BR et al (2006) A Fourier transform infrared microspectroscopic imaging investigation into an animal model exhibiting glioblastoma multiforme. Biochim Biophys Acta 1758:900–907
121. Krafft C, Kirsch M, Beleites et al (2007) Methodology for fiber-optic Raman and FTIR imaging of metastates in mouse brain. Anal Bioanal Chem 389:1133–1142
122. Ali K, Lu Y, Christensen C et al (2008) Fourier transform infrared spectromicroscopy and hierarchical cluster analysis of human meningiomas. Int J Mol Med 21:297–301

123. Beljebbar A, Amharref N, Leveques A et al (2008) Modeling and quantifying biochemical changes in C6 tumor gliomas by Fourier transform infrared imaging. Anal Chem 80:8406–8415
124. Sobottka SB, Geiger KD, Salzer R et al (2009) Suitability of infrared spectroscopic imaging as an intraoperative tool in cerebral glioma surgery. Anal Bioanal Chem 393:187–195
125. Beljebbar A, Dukic S, Amharref N et al (2010) Screening of biochemical/histological changes associated to C6 glioma tumor development by FTIR/PCA imaging. Analyst 135:1090–1097
126. Choo LP, Wetzel DL, Halliday WC et al (1996) In situ characterization of β-amyloid in Alzheimer's diseased tissue by synchrotron Fourier transform infrared microspectrometry. Biophys J 71:1672–1679
127. Miller LM, Wang Q, Telivala TP et al (2006) Synchrotron-based infrared and X-ray imaging shows focalized accumulation of Cu and Zn co-localized with β-amyloid deposits in Alzheimer's disease. J Struct Biol 155:30–37
128. Leskovjan AC, Lanzirotti A, Miller LM (2009) Amyloid plaques in PSAPP mice bind less metal than plaques in human Azheimer's disease. NeuroImage 47:1215–1220
129. Leskovjan AC, Kretlow A, Miller LM (2010) Fourier transform infrared imaging showing reduced unsaturated lipid content in the hippocampus of a mouse model of Alzheimer's disease. Anal Chem 82:2711–2716
130. Gallant M, Rak M, Szeghalmi A et al (2006) Focally elevated creatine detected in amyloid precursor protein (APP) transgenic mice and Alzheimer disease brain tissue. J Biol Chem 281:5–8
131. Kuzyk A, Kastyak M, Agrawal V et al (2010) Association among amyloid plaque, lipid, and creatine in hippocampus of TgCRND8 mouse model for Alzheimer disease. J Biol Chem 285:31202–31207
132. Rak M, Del Bigio MR, Mai S et al (2007) Dense-core and diffuse Aβ plaques in TgCRND8 mice studied with synchrotron FTIR microspectroscopy. Biopolymers 87:207–217
133. Kastyak-Ibrahim MZ, Nasse MJ, Rak M et al (2012) Biochemical label-free tissue imaging with subcellular-resolution synchrotron FTIR with focal array detector. NeuroImage 60:376–383
134. Chwiej J, Dulinska J, Janeczko K et al (2010) Synchrotron FTIR micro-spectroscopy study of the rat hippocampal formation after pilocarpine-evoked seizures. J Chem Neuroanat 40:140–147
135. Dulinska J, Setkowicz Z, Janeczko K et al (2012) Synchrotron radiation Fourier-transform infrared and Raman microspectroscopy study showing an increased frequency of creatine inclusions in the rat hippocampal formation following pilocarpine-evoked seizures. Anal Bioanal Chem 402:2267–2274
136. Kastyak MZ, Szczerbowska-Boruchowska M, Adamek D et al (2010) Pigmented creatine deposits in amyotrophic lateral sclerosis nervous system tissues identified by synchrotron Fourier transform infrared microspectroscopy and X-ray fluorescence spectromicroscopy. Neuroscience 166:1119–1128
137. Szczebowska-Boruchowska M, Dumas P, Kastyak MZ et al (2007) Biomolecular investigation of human substantia nigra in Parkinson's disease by synchrotron radiation Fourier transform infrared microspectroscopy. Arch Biochem Biophys 459:241–248
138. Choo LP, Jackson M, Halliday WC (1993) Infrared spectroscopic characterization of multiple sclerosis plaques in the human central nervous system. Biochim Biophys Acta 1182:333–337
139. LeVine SM, Wetzel DL (1998) Chemical analysis of multiple lesions by FT-IR microspectroscopy. Free Rad Biol Med 25:33–41
140. Heraud P, Caine S, Campanale N et al (2010) Early detection of the chemical changes occurring during the induction and prevention of autoimmune-mediated demyelination detected by FT-IR imaging. NeuroImage 49:1180–1189

141. Pan KM, Baldwin M, Nguyen et al (1993) Conversion of alpha helices into beta-sheets features in the formation of the scrapie prion proteins. Proc Natl Acad Sci U S A 90: 10962–10966
142. Kneipp J, Beekes M, Lasch P et al (2002) Molecular changes of preclinical scrapie can be detected by infrared spectroscopy. J Neurosci 22:2989–2997
143. Kneipp J, Miller LM, Joncic M et al (2003) In situ identification of protein structural changes in prion-infected tissue. Biochim Biophys Acta 1639:152–158
144. Kretlow A, Wang Q, Kneipp J et al (2006) FTIR-microspectroscopy of prion-infected nervous tissue. Biochom Biophys Acta 1758:948–959
145. Liu KZ, Jackson M, Sowa MG et al (1996) Modification of the extracellular matrix following myocardial infarction monitored by FTIR spectroscopy. Biochim Biophys Acta 1315: 73–77
146. Liu KZ, Dixon IMC, Mantsch HH (1999) Distribution of collagen deposition in cardiomyopathic hamster hearts determined by infrared microscopy. Cardiovasc Pathol 8:41–47
147. Gough KM, Zelinski D, Wiens R et al (2003) Fourier transform infrared evaluation of microscopic scarring in the cardiomyopathic heart: effect of chronic AT_1 suppression. Anal Biochem 316:232–242
148. Cheheltani R, Rosano JM, Wang B et al (2012) Fourier transform infrared spectroscopic imaging of cardiac tissue to detect collagen deposition after myocardial infarction. J Biomed Opt 17:056014-1–056014-9
149. Wang Q, Sanad W, Miller LM et al (2005) Infrared imaging of compositional changes in inflammatory cardiomyopathy. Vib Spectrosc 38:217–222
150. Toyran N, Lasch P, Naumann D et al (2006) Early alternations in myocardia and vessels of the diabetic rat heart: an FTIR microspectroscopic study. Biochem J 397:427–436
151. Zohdi V, Wood BR, Pearson JT et al (2012) Evidence of altered biochemical composition in the hearts of adult intrauterine growth-restricted rats. Eur J Nutr. doi:10.1007/s00394-012-0381-x
152. Chiriboga L, Yee H, Diem M (2000) Infrared spectroscopy of human cells and tissue. Part VI: a comparative study of histopathology and infrared microspectroscopy of normal, cirrhotic, and cancerous liver tissue. App Spectrosc 54:1–8
153. Chiriboga L, Yee H, Diem M (2000) Infrared spectroscopy of human cells and tissue. Part VII: FT-IR microspectroscopy of DNase- and RNase-treated normal, cirrhotic, and neoplastic liver tissue. App Spectrosc 54:480–485
154. Liu KZ, Man A, Shaw RA et al (2006) Molecular determination of liver fibrosis by synchrotron infrared microspectroscopy. Biochim Biophys Acta 1758:960–967
155. Le Naour F, Bralet MP, Debois D et al (2009) Chemical imaging on liver steatosis using synchrotron infrared and ToF-SIMS microspectroscopies. PLoS ONE 4:e7408
156. Gautam R, Chandrasekar B, Deobagkar-Lele M et al (2012) Identification of early biomarkers during acetaminophen-induced hepatotoxicity by Fourier transform infrared microspectroscopy. PLoS ONE 7:e45521

Chapter 16
Fourier Transform Infrared (Ft-Ir) Spectroscopic Imaging for Solid Tumor Histopathology

Sreeradha Biswas, Michael J. Walsh and Rohit Bhargava

Abstract Fourier transform infrared (FT-IR) spectroscopic imaging has shown great promise in becoming a powerful tool in cytology and histopathology. Applications for cancer diagnoses in solid tumors are especially attractive as samples are spatially complex and involve myriad molecular changes whereas there are many shortcomings in current clinical practice that can be addressed. Here we review the current state of the art in applying FT-IR imaging for analyzing solid tumors. We focus on instrumentation that is relatively new, emerging fundamental understanding gained by new theoretical advances, data analysis and selected, illustrative applications in cancer histopathology.

Keywords Fourier transform infrared imaging • Theory • Instrumentation • Lasers • Discrete frequency • Tissue imaging • Cancer pathology • Metrics • Bayesian classification • Statistical pattern recognition • Prostate • Breast • Histopathology

16.1 Introduction

Fourier transform infrared (FT-IR) spectroscopic imaging [1] is rapidly emerging as a powerful technique for studying complex biological systems. Advances in instrumentation in combination with rapid development of computational algorithms has led to integration of high throughput data acquisition, robust classification and fast validation processes to constitute practical protocols that can be used in histopathology [2] and other biomedical research [3, 4]. Though the use of IR spectroscopy[1] [5, 6] for biomedical applications has had a long history [7], the last decade

[1] Many issues are summarized in [6].

R. Bhargava (✉)
Departments of Bioengineering, Mechanical Science and Engineering, Electrical and Computer Engineering, Chemical and Biomolecular Engineering and Chemistry, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA
e-mail: rxb@illinois.edu

S. Biswas
Biophysics program, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA

M. J. Walsh
Beckman Institute for Advanced Science and Technology, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA

M. Baranska (ed.), *Optical Spectroscopy and Computational Methods in Biology and Medicine,* Challenges and Advances in Computational Chemistry and Physics 14,
DOI 10.1007/978-94-007-7832-0_16, © Springer Science+Business Media Dordrecht 2014

has seen a fundamental change in using spectroscopy for cancer pathology [8]. Most practitioners now recognize that imaging is absolutely essential for complex tissues, imaging technology has vastly improved and methods to analyze data [9] are better understood and better applied. Exciting advances in all spheres of the development of this technology are taking place in a renaissance of sorts [1]. A particular area of focus has been biomedical applications [10] and, especially, in disease diagnosis. Here, we review the needs in pathology, recent developments in various aspects of the technology and present a brief description of some applications.

16.2 Current Techniques in Clinical Histopathology

16.2.1 General Methods and Techniques—A Brief Overview

Histopathology refers to the microscopic examination of the different cell types in a tissue with the aim of diagnosing signs of disease. One of the most important areas in which this technique may have impact is in cancer diagnosis, grading and guidance of therapy. The procedure employed today for determining whether a tissue has malignant disease or not after suspicious screening is as follows: a biopsy is employed first to obtain a sample from the suspicious area from the affected tissue, it is then fixed, for example—with formalin, to prevent degradation of the tissue and maintain structural integrity during subsequent processing. Next, it is chemically processed and embedded in paraffin wax, such that well-preserved two-dimensional sections, 3–7 μm thick, can be sliced from it with the help of a microtome. These sections are then most commonly floated onto glass slides; the slides are washed and stained, following a standard protocol. Native tissue does not have much contrast in brightfield microscopy. Hence, a stain is used to allow a pathologist to visualize structure. The most common stain used is a combination of hematoxylin and eosin (H&E). Hematoxylin stains the nuclei blue, whereas eosin stains cytoplasm and the extracellular matrix pink. Other types of stains commonly used include safranin, Oil Red O, congo red, silver salts and other artificial dyes. The purpose of staining is to provide contrast between the various cell types in the tissue and counterstains can also be used further to make visualization clearer by producing additional contrast. Pathologists are trained to distinguish the different cell types by observing stained tissues using a microscope, as shown in Fig. 16.1.

16.2.2 Manual Diagnostic Techniques

Morphology of individual cells as well as their patterns to form functional units such as glands, shapes of both cells and glands, numbers and (distorted) organizations of cells all provide useful evidence to differentiate between normal and malignant tissues. A pathologist typically examines the tissue at different magnifications and determines by manual decision-making whether the tissue can be classified into a specific grade. As may be imagined, there are several shortcomings to this manual

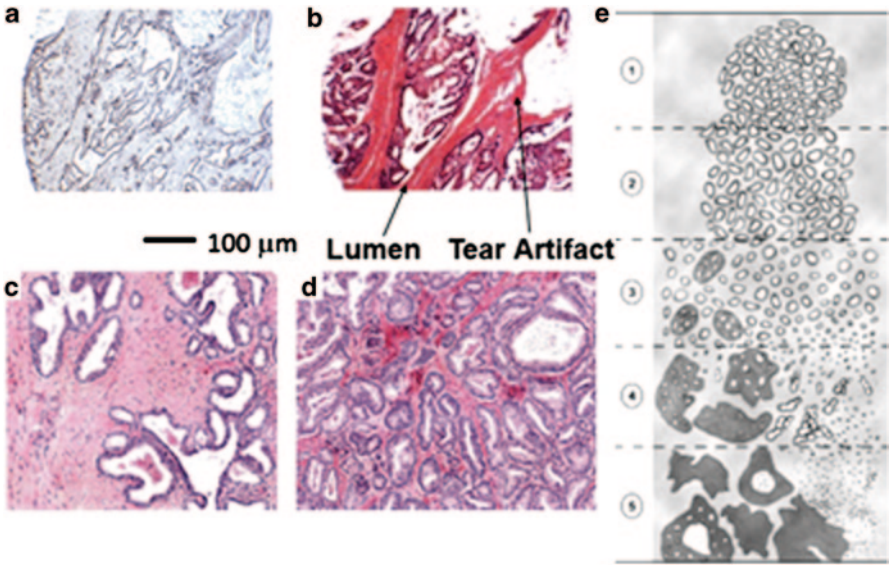


Fig. 16.1 Prostate histology and carcinoma grading. **a** Unstained tissue has little contrast, requiring staining for brightfield microscopy (**b**). Epithelial cells define empty regions (*lumen*) that are well-differentiated in normal (**c**), and poorly in cancerous tissue (**d**). If lumen are nearly regular and well-spaced, tissue is benign. Increasing disruption of this structure leads to increasing grade of tumors, from 1 to 5, on the Gleason Scale (**e**). (Based on figure from ref. [2])

and subjective process. Efficiency, cost and accuracy can be improved. Most importantly, the methods for most tissues are at least several decades old and have not kept pace with rapid scientific advances that have enabled modern science in a number of fields. The same morphologic measures also help to diagnose the stage of disease or other specific characteristics that provide critical input for determining appropriate therapy. It is well recognized that many diagnoses in tissues are not absolutely accurate and differences in diagnoses between different pathologists as well as mistakes are often made. Hence, the entire process of diagnosing disease and recommending therapy can be updated with other information. Lately, this process has been extensively supplemented with molecularly-specific stains. In some cases, the staining characteristics directly inform molecular medicine to be administered to the patient. For example, estrogen receptor positive (ER+) breast tumors often respond to tamoxifen. In general, modern trends in pathology include digital recording and analysis as well as increased integration of molecular content. These two trends are highly resonant with the goals of using IR imaging for histopathology.

16.2.3 Molecular Histopathology Techniques

At times, definitive diagnoses might not be possible and research questions may not be answerable by analyzing generic stains. The reasons can be several—the

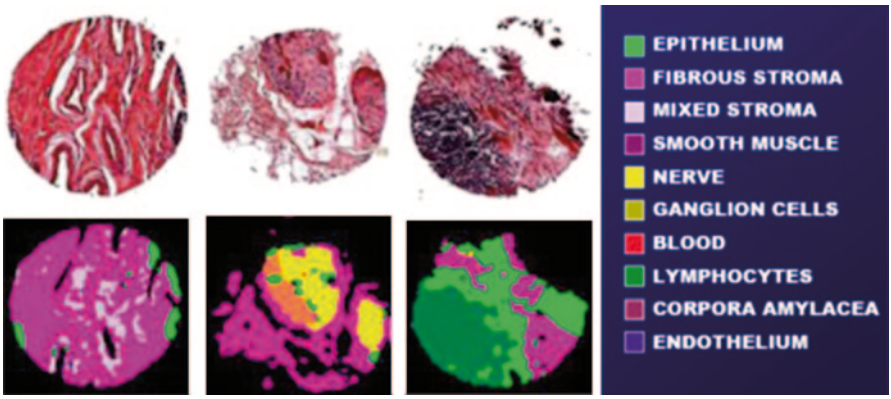


Fig. 16.2 Segmentation of prostate tissue into constituent cell types using IR spectral data at every pixel. Pixel-by-pixel enumeration, quantitative analysis and segmentation are possible, compared to human ability. Hence, quantitative analyses of the abundance, position and chemistry of different cell types is possible. (Based on figure from ref. [11])

biopsied tissue available for examination might be too small which makes diagnosis difficult or the case might be complicated—it might not be possible to confirm the diagnoses with confidence or molecular information may be required in addition to the morphologic information. In these cases, molecular histopathology has been shown to add significant additional information that can aid the pathologists in making a definitive diagnosis. The most important molecular technique employed by pathologists is immunohistochemical (IHC) staining. The principle behind IHC staining is the detection of a particular antigen in a tissue-section using an antibody which specifically binds to that antigen. IHC stains can thus act as biomarkers and they can be used in the detection of particular cell types with high specificity and accuracy. Some common immunohistochemical stains are ACTH—Adrenocorticotropic hormone: for classifying pituitary adenomas, Calponin: for detecting smooth muscle cells, myopithelial cells and certain stromal myofibroblasts, Smooth Muscle ActinSMA: for visualising smooth muscular tissue along with myoepithelium, and myofibroblasts, used in distinguishing between leiomyosarcoma from rhabdomyosarcoma,Vimentin: for staining fibroblasts and myofibroblasts, used in diagnosing melanomas and sarcomas. Unfortunately, the disadvantage is that correct diagnosis by IHC staining is dependent on the accuracy of the staining procedure as well as the skill of the pathologist; it is a highly sensitive process and hence can lead to variable results with even slight modifications. Hence, it is not normally used as a primary diagnostic standard but mainly, for conforming decisions in complicated cases or for molecular characteristics to guide therapy. IR imaging can fill this void without the use of stains but by using computation to provide the contrast, as illustrated in Fig. 16.2 [11]. In this paradigm, color coded images replace conventional stained images. The information in color code is the same as would be determined by an expert pathologist after review of the conventionally-stained images, but the recognition is accomplished via computer algorithms.

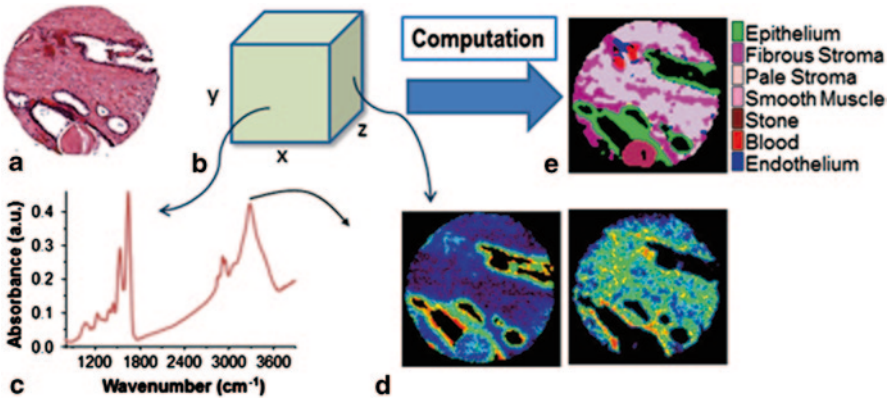


Fig. 16.3 **a** Conventional imaging in pathology requires dyes and a human to recognize cells. In chemical imaging data cubes (**b**), both a spectrum at any pixel (**c**), and the spatial distribution of any spectral feature can be seen. e.g. in **d** nucleic acids (*left*, at ~1,080 cm^{-1}), and collagen specific (*right*, at ~1,245 cm^{-1}) Computational tools can then convert chemical imaging data to knowledge used in pathology (**e**). (Based on figure from ref. [2])

16.3 IR Spectroscopic Imaging for Automated Histopathology

An overview of the application of IR imaging [12] for histopathology [13] is shown in Fig. 16.3. As opposed to stained images in which 3 colors of the visible spectrum are acquired, much finer spectral data is acquired in IR imaging, resulting in a data set in which a spectrum at every pixel can be used to make decisions about the content and identity of the biological material at that pixel using computer algorithms. Easy visualization of the information contained within the image can be provided by the color coded images, by summary statistics or even by converting the data into binary answers or suggestions. Probabilistic information is enabled by the extensive underlying information and databases underpining the diagnosis and computer algorithms used present extensive opportunities for integrating other forms of information or interactive decision-making. The use of IR spectroscopic imaging for histopathology, thus, is not simply an attempt to reproduce what is already accomplished in biomedical analysis but can substantially extend present capabilities in innovative and useful ways.

16.3.1 Data Acquisition

The first step towards automated histopathology is to acquire IR spectral images. Perhaps the two most critical hardware components are the source and the detector. While the choice of detectors primarily determines the type of data acquisition, speed, scanning mode and optical setup, the evolution of sources similarly influences the process. While several formats and types of detectors are available, the choice of sources is still fairly restricted. The evolution of technology for both these components will likely influence the recording of data for many years.

16.3.1.1 IR Mapping

For several decades now, IR spectrometers have used thermal sources coupled to microscopes equipped with single-element detectors for acquiring point spectra. Spectra recorded could be assigned to pixels in a sequential fashion with the help of raster scanning, to create a point-by-point mapped image. There were two big disadvantages to this process. First, the signal to noise ratio (SNR) was poor because benchtop thermal sources are of relatively low power. Spatial resolution was poor as well due to using IR opaque apertures to define the spatial resolution. In general, the IR absorbance images formed lacked clarity and accuracy, especially when compared to optical microscopy. Second, due to raster scanning a single pixel, data acquisition was an extremely slow process and scanning tissue sections of reasonable size was not feasible.

Recent advances in scanning technology, noise filtering and better hardware have resulted in some impressive advances. New products are still being introduced with single element detectors that provide exceptional performance. One advance has been to use smaller AC-coupled detectors, in which low noise can be further rejected by electronic filtering. The SNR of such systems can be exceptionally high and can provide exceptional spectra even from small aperture sizes due to better engineered systems that are dedicated to single element data recording. A notable exception has always been the use of synchrotron sources, which allowed for small aperture sizes and exceptional quality data [14–16]. This coupling essentially provides a large signal [17]. While there is no doubt that the large signal improves throughput and is very effective for small aperture sizes, simply using a brighter source is not a guarantee of the best performance. For example, the brighter source may require faster mirror scanning, leading to larger Fourier frequencies and other sources of noise that may overwhelm detector noise. Shot noise in the source may also start to become apparent in some cases on using a brighter source. Hence, we posit that any instrument that is designed around a specific source or detector carefully also consider data acquisition. While this is a fine point and a simplification is often made that using a "better" source or detector will improve the measurements, it is truly the integrated system in which every matched part is optimally designed, that will provide the best performance. It turns out that the flux provided by a synchrotron does not lead to unstable recording or encroachment into regimes where the intensity is a problem, emerging laser sources must be cautiously integrated.

Synchrotron sources are of significant importance because synchrotron beams can be collimated very efficiently and can form bright, very high intensity sources that are almost point-like. Covering point areas of the sample and, as with globar sources, single-element detectors used for acquiring point spectra by raster scanning can create a mapped image. This makes them an ideal source to be used to validate fundamental theoretical advances in IR microscopy. In some ways, the synchrotron is an ideal source for IR microscopy. While it is obvious that the cost and availability of such systems cannot compete with the globar, synchrotron-based systems nevertheless represent a useful avenue for proof-of-principle studies and provide a performance yardstick for all of IR microscopy. Even under these conditions, there is one disadvantage

for cancer histopathology: data acquisition is time-consuming and imaging reasonably-sized areas is difficult simply because of the need to scan the beam or the samples. Given the high SNR of individual spectra obtained with the widely-available, cheaper detectors and, however, we anticipate that efforts will be made to further reduce scanning time and make the point scanning approach even faster.

16.3.1.2 IR Imaging

As opposed to using a single element detector for microscopy, a number of approaches have been proposed for using an array detector [18]. Array detection offers a multichannel detection advantage, considerable speeding up data acquisition [19, 20]. For example, acquisition of thousands of spectra is now possible in nearly the same time period as a single spectrum in point microscopy. The acquired data, however, are not of as high an SNR as their single element counterparts. A thermal source is still widely used as the IR source but has to illuminate a large sample compared to that of the single-element detector-equipped systems. Multi-element detectors or focal plane array (FPA) detectors often allow measurements of a large number of pixels (typically 16,384 in the most common implementation today). Since the array detector contains miniaturized individual detectors, the readout electronics of each detector also have to be miniaturized. This reduction in size reduces the quality of the signal readout and results in a higher noise level in every pixel compared to the single element detector that can have macroscopic electronics. The loss in SNR can somewhat be compensated by scanning faster and signal averaging extensively [21]. Compared to the early detector systems, frame rates of over 1.5 kHz now permit data to be rapidly and simultaneously recorded from all pixels in the FPA. The analog to digital converter of FPA pixels are often filled by microsecond integration time, though the frame rates are in the millisecond regime [22, 23]. Hence, there is still considerable potential in making faster detectors for IR imaging. We anticipate that at least 20 kHz speeds can be sustained without any degradation in performance [24]. Advances in readout architecture are also on the horizon. The conversion from analog to digital readout from the chip will also lead to lower noise, perhaps by more than an order of magnitude. Hence, advances in detector technology alone can make scanning two orders of magnitude faster. The very rapid sampling will considerable enhance studying biological processes, including time resolved observations like biochemical kinetics of living cells. The faster detectors will also enable the acquisition of a large number of pixels rapidly. This aspect has become more critical with the advent of high-definition IR imaging.

16.3.1.3 High Definition IR Imaging

An exciting new development is the realization of high-definition IR imaging [25–27]. In an effort to take diffraction limited images at a high SNR at every pixel, it was long believed that the optimal size of the pixel at the sample plane was given

by the Rayleigh criterion; namely, it was believed that the smallest pixel size at the sample focal plane should be the diffraction limited spot size and no further improvement in image quality was possible. Recently, it was realized that this was not the case. The appropriate pixel size can be understood in two contexts, both of which are related. First, it is trivial to observe that there is structure in the point spread function (PSF) of any focusing optic. To access this structure, clearly the pixel size must be smaller. A suggestion of 5 pixels to sample the entire PSF, implies that the effective pixel size be approximately five-fold smaller than the Rayleigh prescribed spot size. This leads directly to a case where the pixels size is approximately 0.5 micron, at the short end of the mid-IR wavelengths. Since this implies that a 100-fold smaller pixel is optimal compared to most commercial instruments, a much brighter source was needed. The concepts of proper sampling and smaller pixels was achieved using a high brightness synchrotron source coupled to a FPA detector [25]. The major problem in realizing this concept was that the synchrotron source could illuminate only a small sample area due to its intrinsically smaller size; hence, either a small number of pixels of the FPA could be used for detection, or the number of scans had to be significantly increased in order to detect the very low signal per pixel in a diffused beam. The IRENI (IR Environmental Imaging) at the Synchrotron Radiation Center in Madison, WI was the world's first IR beamline to be able to overcome this problem by combining a multiple synchrotron beam source, along with an FPA detector [28]. This and other systems [29–32] utilize the advantages of both large-area illumination from a high brightness, stable, broadband synchrotron source and resulting high SNR and spatial resolution along with the high speed and large area scanning ability of multiple element FPA detectors. Two additional changes were made in order to achieve the optimal imaging for a given wavelength. First, a high numerical aperture (NA) objective (according to the Rayleigh criterion) and, second, pixel magnification small enough to be able to accurately scan minute details of the sample.

According to the Rayleigh criterion, the spatial resolution of a microscope is given by $\frac{1.22\lambda}{2NA}$, where NA is given by $n \sin \theta$, where n is the refractive index of the medium and θ is the maximum subtended angle of the objective. As mentioned previously, commercially-available FT-IR imaging spectrometers are designed for a pixel size of $\sim 5 \times 5$ μm to correspond to the mid-band of the mid-IR wavelengths and a NA of ~0.5. The IRENI source on the other hand, used an objective of NA—0.65, thus increasing its spatial resolution as well. The combination of 12 bright synchrotron radiation sources in the IRENI set up, is strong enough to enable the increase of the magnification of the lens to 74x such that, the effective sample-plane pixel size can be decreased to 0.54×0.54 μm^2. Thus both the criteria are satisfied and this leads to a significant increase in spatial resolution and quality of the image formed as shown in Fig. 16.4. As a first demonstrative example of the new capabilities, we used it to image, identify and characterize small cell types such as those in the intralobular stroma in the human breast and the basement membrane in the human prostate tissue with unprecedented detail.

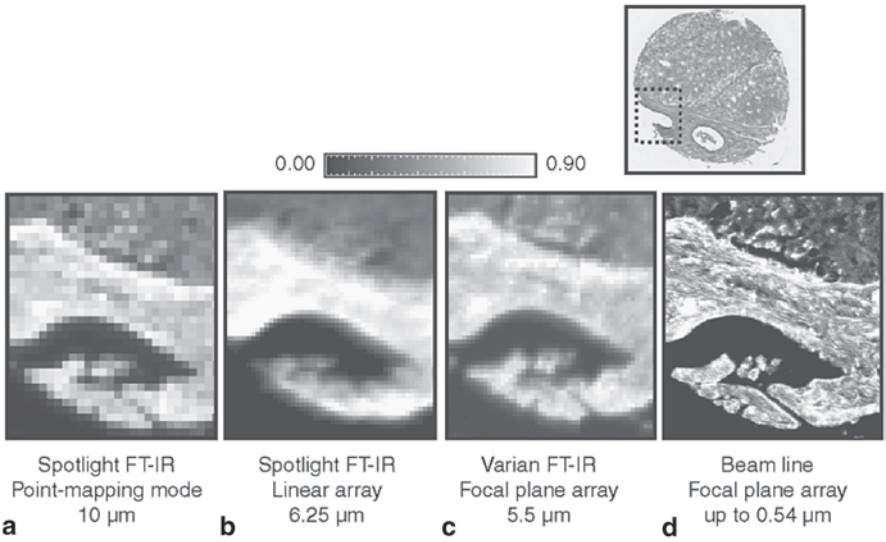


Fig. 16.4 Imaging data from **a** 10 µm aperture in point mapping mode. **b** 6.25 µm pixel size at the sample plane in a multichannel linear detector (scanning mode). **c** 5.5 µm pixel size using a 2D focal plane array detector. **d** 0.54 µm pixel size with the synchrotron beam. (Figure based on ref. [25])

Another, more rigorous approach to understanding the quality of image formation was proposed using a Fourier optics approach [27]. A first principles understanding of the achievable quality as a function of instrument parameters was proposed, primarily to affect improved instrument design. A model for light propagation through an IR spectroscopic imaging system was first developed using scalar wave theory. The model analytically describes the propagation of light along the entire beam path between the source and the detector, while explicitly modeling each component in the beam path. The effect of all optical elements and the sample in the microscope was understood in terms of the accessible spatial frequencies in the Fourier optics framework. Subsequently, simulations were conducted to explicitly obtain the image that would be recorded under certain conditions. The modeled and measured images for both low NA and high NA systems, at specific detector pixel spacing agreed remarkably. From this validated model, the optimal pixel size at the sample plane is calculated and shown to be much smaller than that in current mid-infrared microscopy systems but larger than that obtained for the IRENI system. A plot for the optimal pixel size as a function of optics and of wavelength was provided such that any optimization point could be chosen rigorously. A commercial imaging system was then modified and experimental data were presented to validate the developed model. Notably, this demonstration was accomplished on a commercial benchtop IR imaging instrument. Without the use of a synchrotron, acquired data were of high spatial quality but, as expected, of poorer SNR. Signal processing approaches were implemented to improve the spectral SNR. At present, a number of studies have produced data demonstrating the ability to perform

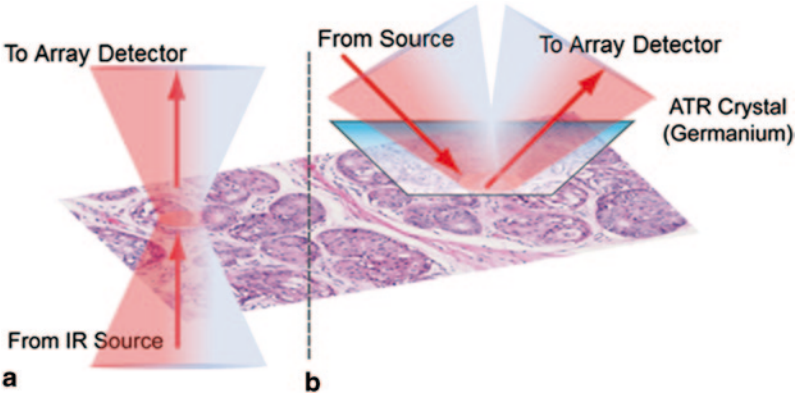


Fig. 16.5 Schematic drawing of **a** transmission-mode microscopy and **b** ATR-mode microscopy for chemical imaging of breast tissue biopsies. In transmission-reflection mode, the ATR crystal is not present and a variety of incidence angles may be used, ranging from 90°, i.e. normal to the surface of the sample and with low NA to grazing angles and very high NA. (Figure based on ref. [33])

high-definition IR imaging in the laboratory using minimally-modified commercial instruments. This is obviously more accessible and cost effective compared to a synchrotron. This dual capability in the community, of rapid imaging using a synchrotron and slower laboratory based systems is likely to prove useful for many practitioners and, doubtless, numerous new studies will appear using the high-definition approach. We anticipate that this optimal configuration will become a standard approach for imaging in the mid-IR.

16.3.2 Sampling Modes for Acquiring IR Imaging Data

While we have considered the design of the microscope and optimal configurations, one of the most critical aspects of conducting IR imaging experiments is the manner in which light interacts with the sample. There are three most commonly used modes for acquiring IR data. These are in transmission, transmission-reflection and attenuated total reflectance (ATR) mode.

16.3.2.1 Transmission Mode

The transmission sampling mode (Fig. 16.5a) is the most common IR imaging configuration and has been reported in a number of studies on human tissues [33]. In this technique, films, cells or tissue sections that are a few micrometers thick are either placed on a single substrate or sandwiched between substrates transmitting IR radiation. Light simply passes through the sample and can be introduced on one side and collected on the other. The geometry is rather unremarkable, except

when placing the sample and using a second substrate to sandwich the sample. The sample, when placed on a single substrate, should be placed such that the substrate is between the condenser and sample. The sample is exposed to the objective such that the focal spot can be formed through air and onto the sample. When a substrate at the top is used, it has been suggested that focusing may be more difficult. Hence, the top substrate should be avoided for both this reason and for the possibility of causing a cavity induced fringing in the data. By using a novel approach, it was recently shown that the top substrate can be structured to actually improve data quality. Kazarian et al. [34, 35] used a hemi-spherical top substrate and demonstrated that the substrate can act as an additional lens to improve magnification in the optical path. The resulting magnification brings the pixel size closer to the optimal sampling and, essentially, results in a high-definition type configuration.

16.3.2.2 Transmission-reflection Mode

The transmission-reflection mode (Fig. 16.5b) is attractive as transmission substrates for IR spectroscopy need to be made of salt crystals and are, consequently, expensive, fragile and often not readily available. The sample usually does not contain an upper substrate and the bottom supporting substrate contains a reflective (often metallic) coating. The sample preparation is relatively straightforward and rather similar to that for transmission mode. Given the double pass of the beam through the sample, however, the optimal thickness of the sample is usually smaller than that for transmission mode. The objective doubles as a condenser in this case. At times, half the objective is employed to illuminate the sample and half to acquire data. This results in an asymmetric resolution capability and different quality in the two orthogonal spatial directions. In practical terms, this shortcoming does not result in an appreciably distorted image, partly because most studies have only examined the larger pixel case. A more serious effect of the sampling geometry is the effect on acquired data. The sample acts as a resonant cavity, with light reflected from the interfaces undergoing multiple reflections and demonstrating the cumulative effect of reflections, scattering and coherent addition. The resulting data are generally difficult to interpret unless a detailed analysis capturing the physics of the process is employed. Consequently, there has been much discussion about the utility of this sampling mode and its use for biomedical samples [36–41].

16.3.2.3 Attenuated Total Reflectance (ATR) Mode

As shown in Fig. 16.5c, an ATR crystal is used either as a substrate-objective or as an overlaid Solid Immersion Lens (SIL) for imaging in this mode. The lens is made of a material of high refractive index, for example, commonly using Germanium, Diamond or Zinc Selenide. In all cases, light is introduced at an angle into the higher refractive index crystal and transmits to the the lower refractive index sample interface where it may suffer total internal reflection if the reflction condition is

matched. An exponentially decreasing evanescent wave is set up at the interface, which penetrates into the sample. The depth of penetration of the evanescent wave depends on the incident angle, refractive indices of the crystal and the sample, and the radiation wavelength. In this case, therefore, the sample does not directly interact with the IR beam, rather it's layers closest to the crystal surface interact with the decaying evanescent wave in a depth weighted manner (19). For commonly used ATR lenses, the depth of penetration is typically of the order of the wavelength of light in the IR. These properties can be used to record data from both thick samples as well as from depth-resolved areas of a sample.

Some of the drawbacks of imaging by transmission or transflectance modes have been solved using the ATR mode: first, the problem of water absorption can be mitigated due to the shallow penetration depth and live, fresh or frozen tissues can be imaged. Second, the NA of the transmission mode characteristically ranges from 0.4 to 0.8. Hence, the smallest beam spot size is approximately equal to the wavelength. Given the high refractive index of the SIL, upto a 4-fold smaller beam spot size can be attained, making it possible to study small cells like endothelial cells or myoepithelial cells, which may be important for diagnosing cancer. As a demonstrative example, using ATR FT-IR imaging (spatial resolution better than the pixel size of 1.56×1.56 μm), finer structure in breast tissue could be imaged that was not visible in the other sampling modes as shown in Fig. 16.6. One of the major advantages of this mode is that sample preparation can be minimal, because processes like sectioning, fixation, or staining can be avoided. ATR imaging hence can be used in some biomedical applications where the Transmission mode would fail to provide sufficient light intensity for measurements, for cases where finer structure needs to be imaged and for those situations in which preparing thin sections of samples is difficult. Due to sub-wavelength coupling, the approach may also allow for the analysis samples with strongly scattering profiles. For example, the analysis of kidney stones without being affected by strong scattering from specular reflection, Christiansen effect and Restsrahlen effects unlike the transmission mode [42].

16.3.3 Fundamental Theoretical Understanding of FT-IR Microspectroscopy and Imaging

One of the most exciting developments in recent years has been the evolution of theoretical understanding in microscopy and image formation in the mid-IR [43, 44]. The most common theoretical treatment for IR microscopy had been to utilize ray or geometric optics. While these approaches were useful in designing instruments and generally predicting the flux and spot size at the focal plane, they are insufficient in describing optical phenomena when the wavelength is approximately equal to sample dimensions. Due to strong absorption, samples are typically microns thick in the mid-IR. Many biological samples are heterogeneous on the micron scale as well. Thus, there are strong effects in the recorded spectrum that arise from the coupling of micron-wavelength light with micron-dimension samples. The effects include

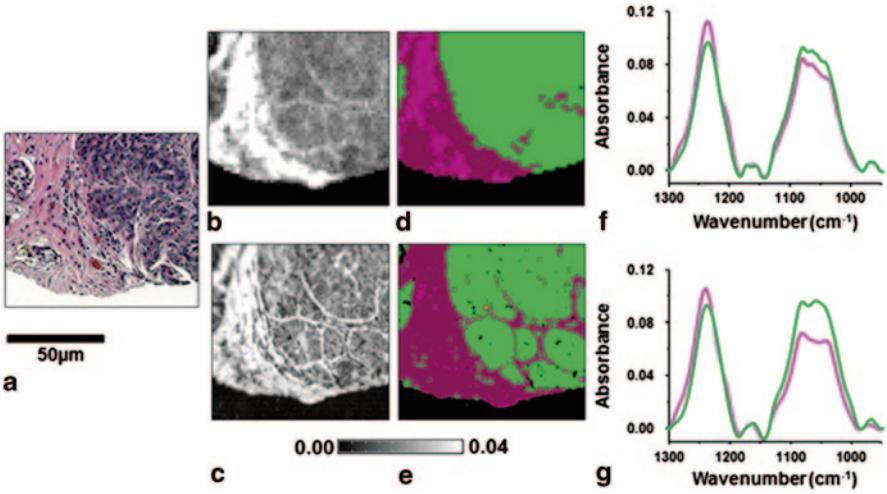


Fig. 16.6 **a** Optical microscopy image of H&E-stained normal breast tissue terminal ductal lobular unit (TDLU). **b** Transmission- and **c** ATR imaging data were recorded on a serial section of the same TDLU from which the absorbance associated with collagen absorption at 1,255 cm^{-1} displays contrast. A cell type classifier was employed to **d** transmission microscopy and **e** ATR microscopy data in which epithelial cells are identified in green and stromal cells in magenta. Average spectra were extracted from the regions associated with the two cell types from the **f** transmission (196 pixels for stroma and 260 pixels for epithelial cells) and **g** ATR imaging data (1,271 pixels for stroma and 6,339 pixels for epithelial cells) to identify chemical differences between the cell types. (Figure from ref. [33])

both macroscopic effects such as refraction and reflection as well as progressively structure-dependent effects including scattering and changes in the interaction volume of the sample. Light interaction consists of the more obvious absorbance arising from the imaginary part of the refractive index and others that result in no loss of intensity but changes in phase (direction) arising from the real part of the refractive index. Since the real and imaginary parts of the refractive index are coupled, optical phenomena primarily associated with one or the other affect images with either contrast mechanism.

Electromagnetic theory can provide a detailed analysis of the optical physics involved in IR imaging and has been recently used to describe the image formation process. Application to increasingly complex shapes and geometries has been provided in a series of papers. Typically, in this approach, electric and magnetic fields associated with a monochromatic plane wave are considered to be the building blocks in the theoretical treatment. The system is assumed to be linear; hence, summation of individual plane waves can be applied to reconstruct the field as well as all linear transformations of the electomagnetic field. For each plane wave constituting the electromagnetic field, its quantitative contribution is given by the amplitude and its relative position is given by the phase. By propagating a collection of plane waves through the instrument and interactions with the sample, precise

predictions of acquired data can be made. While the propagation of light through optical components can be modeled using methods available from other spectral regions, the interaction of light with the sample is an area of continuing development. While specific geometries such as spheres, cylinders or slabs can be modeled, the general case remains to be shown in an efficient manner. Nevertheless, the framework presented forms the basis for physics-based modeling of any sample. Though EM theory is rigorous and can provide the most detailed picture of light-matter interactions, its implementation and interpretation of results are both significantly complex. Recently, a conceptually simpler approach has been proposed to understand the optics of microscopy using scalar wave theory [24]. In this approach, the major difference is that the wave propagation is simplified and computations can potentially be faster. We invoked the same structural model of the sample as in the complete EM treatment, in terms of decomposition into Fourier frequencies. Together, the two approaches should be useful in describing light matter interactions in a number of sample geometries and configurations [45, 46].

16.3.4 Applications of IR Imaging in Solid Tumor Histopathology and Corresponding Normal Tissues

In this chapter, we do not aim to thoroughly cover every example of application of IR imaging for histopathology. We also do not cover the older work using single elements detectors. Many excellent reviews exist to provide the reader a guidance of the development of the field and a comprehensive view of the variety of possible applications. Instead, we focus here on providing an overview and a flavor of the possibilities using IR imaging. We also point to salient aspects of the process and key features of any diagnostic protocol using illustrative examples.

16.3.4.1 Basic Concept of Using Absorption Spectra for Medical Diagnostics

Biological samples are composed of several complex small molecule and macromolecular assemblies, including proteins, nucleic acids, lipids and carbohydrates, as well as constituent small molecules [47]. Very rarely (with notable exceptions such as collagen) do spectra reflect compositions of any of these constitutents directly. Hence, IR absorption spectra contain this entire diversity, simultaneously making them both very sensitive to all constituents and exceptionally non-specific to any particular molecule. Spectra comprise of some significant, spectral characteristic features, however, and significant “mesoscale” chemistry can be examined as, in general, classes of compounds can be distinguished. For example, Fig. 16.7 demonstrates the spectra from different classes of compounds that make up the bulk of human cells and tissues. Spectra from these four classes are markedly different from each other. At a finer level, spectra from specific molecules of each type may also demonstrate differences, but those will be much finer than the differences be-

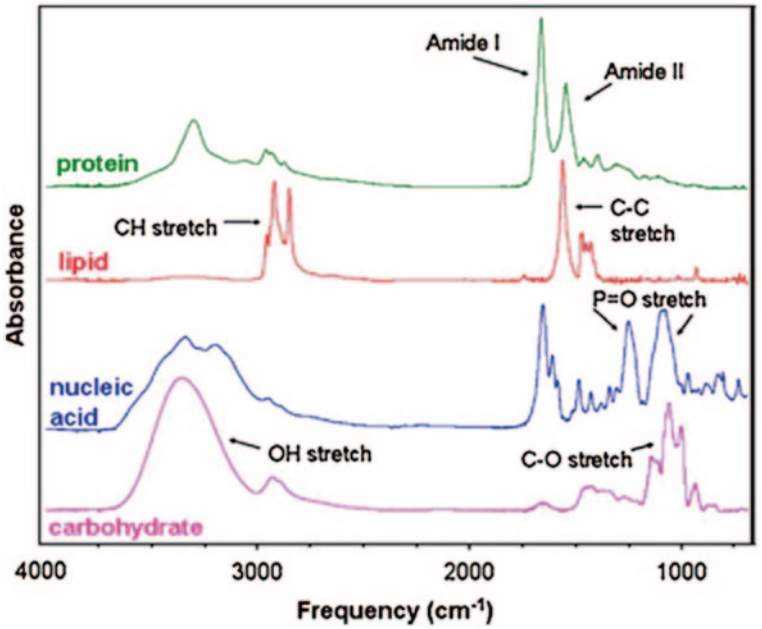


Fig. 16.7 FT- IR absorption spectra of biomolecules: (protein (myoglobin), lipid (dimyristoylphosphatidylcholine, DMPC), nucleic acid (poly-A), and carbohydrate (sucrose)).The absorption peaks for different biomolecules differ due to difference of composition and conformation, thus lending specific, distinguishing characteristic features to each biomolecule. (Reproduced from ref. [47])

tween classes. Typically, it is this combination of spectral features that are observed in most biological samples and the gross distinction between different biological or physiological conditions is significantly influenced by the differences in these classes of compounds. The use of imaging enhances sensitivity to spectra differences by localizing signal. It is estimated that the sensitivity to differences likely lies in the pg/pixel range for modern instruments.

In a typical biological sample, the spectrum is dominated by the the amide I band can be observed at ~1,650 cm^{-1} as a result mainly of the ν(C-O) stretching vibrations of the protein amide bonds; the amide II band appears around 1,550 cm^{-1} and is a result of the δ(N–H) bending and ν(C–N) stretching vibrations of the amide bonds. The ν(C-O) stretching mode of lipids is observed ~1,740 cm^{-1} while the antisymmetric ν_{as} and symmetric ν_s (PO_2^-) stretching mode of nucleic acids and phospholipids are around 1,225 and 1,080 cm^{-1}.Thus, by analyzing the absorption spectra, it is possible to predict composition to a certain extent. Perhaps of greater relevance is that the relative spectral features provide a signature of the composition and physiologic state of the biologic. In tissue, histologic classes such as 'epithelium', 'stroma', 'blood', etc. may contain similar biochemical components (proteins, nucleic acids, carbohydrates and lipids), but their compositions are different;

hence their characteristic spectra are similar but show small differences in some specific features. Chemical concentrations can be measured by recording the spectral absorbance at specific vibrational frequencies, while subtle molecular structural changes are analysed by studying the spectral peak shifts, band shapes, relative intensity changes portrayed in the absorption spectra of the respective classes [48]. In most IR imaging experiments, determinations are made pixel-by-pixel.

An absorption spectrum is obtained for every pixel of a 2D (thin in terms of beam interaction in one dimension compared to the other two dimensions), thus leading to the formation of the so-called "data cube". The acquired data is usually analyzed in two ways: either by obtaining chemical images at every wavenumber or by extracting absorption spectra at every pixel. For applications in histopathology, both methods have advantages but spatial information in imaging is typically of more interest and spectral information is usually compressed in some manner to be spatially appreciated. The most common is the use of color coding[2], in which a color coded image demonstrates the presence of multiple cell types in a tissue.

16.3.4.2 Building a Classifier

To relate spectral features to underlying biologic or physiologic assessment, a computational approach typically based on statistical properties of the data is most commonly used. Typically, spectral features from the absorption spectra are fed into pattern recognition algorithms to extract information. An overview of the various steps that may be typically involved in a specific protocol are shown in Fig. 16.8 [3]. The key elements of building the classifier are discussed below.

Data Acquisition While it is obvious that data acquisition is the first step, a critical consideration at this stage is to ensure that the data are representative of the full data set. If this is not ensured, developed methods may be subject to the dual pitfalls of chance and bias, limiting the applicability of any developed protocol. The data set to develop protocols must also be sufficiently large for the results to have required statistical power.

Pre-Processing Pre-processing collectively refers to a set of computational processes that aim to enhance the characteristics of the acquired data, making analysis more accurate. The use of algorithms such as the Minimum Noise Fraction to reduce noise for more accurate analysis, baseline correction to remove effects due to scattering and removal of water vapor signals are all examples of common steps. Appropriate pre-processing plays a very important role in both analyzing and extracting important chemical information from recorded spectra and paves the path to building accurate, fast and robust classifiers.

Classification The goal of classification methods is to categorize data into subsequent groups, which can provide with powerful information about the molecular

[2] Note that some texts refer to this practice as "False Color". This term is misleading as all colors are true. "Color-coded" images or "Representative color" images are probably better monikers.

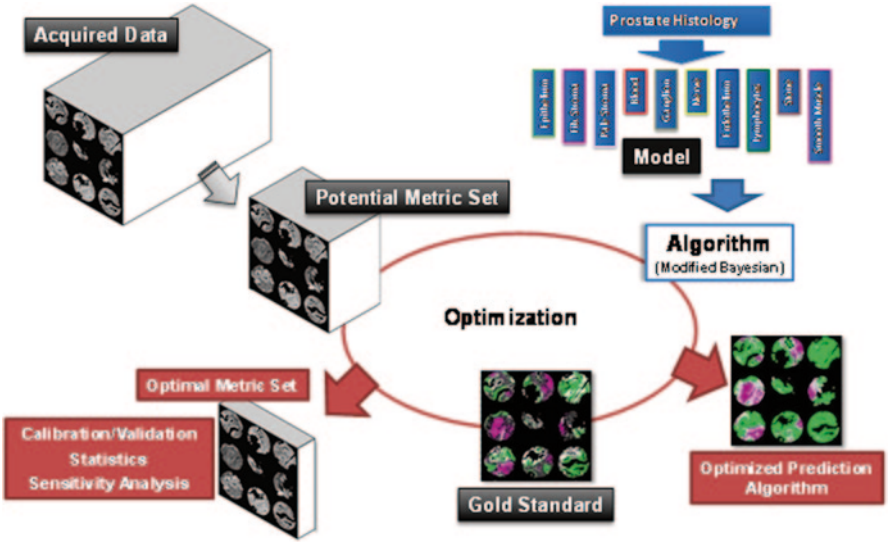


Fig. 16.8 Schematic diagram of the process involved in developing cell-type classifiers using FT-IR imaging. Models are built which use information from the pathologist (cell type, disease state) and are used to create classifiers. Recorded data is turned to a potential set of predictors (metrics). Using a standard pattern recognition algorithm, metrics are chosen and a completely automated and optimized prediction algorithm is developed. Calibration and validation statistics as well as model robustness can also be gauged. (Reproduced from ref. [3])

and/or structural composition of the samples as well as relationship to physiologic or disease processes. Classification techniques can be broadly divided into two types, namely, supervised and unsupervised methods.

Supervised Methods In the use of supervised methods, the final categorization of data is known a priori and accurately marked representative samples from these known classes are available. It is critically important that the size and characteristics of the representative samples should be such that they represent the entire population well. These representative samples constitutes the training data set. A similarly representative data set is separately selected and employed to validate the results obtained from the methods. Appropriately, this data set is termed the validation set.

The development of an analysis protocol using supervised classification can be divided into three parts: first, the training data set is acquired and subjected to a series of pre-processing steps to enhance data quality. Second, the class labels applied to training data are used to identify potentially significant and differentiating features within the spectrum. For example, in imaging human tissues, it is common to differentiate histological classes such as ‘epithelium’, ‘stroma’ and ‘blood’, among others. Spectra from different classes show differences in absorption spectra because of the differences in their chemical compositions. Hence, spectra from each class are analyzed carefully to deduce characteristic differences – spatial, spectral, or both – that highlight differences between specific cell types. There are

two approaches to the analysis of spectral features. In the first, eigenvectors of the entire spectrum can be considered in some transform space (for example, using principal components analysis or PCA). In this approach, all the information contained in the spectrum is utilized, making it very powerful. In the opinion of the authors, this approach also includes all scattering and information dependent on the sample's physical characteristics, making its use for classification highly dependent on the physical characteristics of the data set. In a second approach [48], spectral knowledge and biochemical insight are used to define spectral features that may potentially be used for classification. These are named as 'metrics'. The use of metrics can easily avoid factors other than chemical differences but may miss some features useful for classification. An example of a metric can be the ratio of absorbance of the phosphodiester-specific spectral peak at 1,080 cm^{-1} to the amide II vibration-specific absorbance at 1,545 cm^{-1}. This ratio has shown to be useful in discriminating epithelial cells from stromal cells. Recently, an analysis of variance on IR imaging data sets [49] revealed that the sample morphology effects on spectra may be large compared to the information bearing metrics for some cases. This evidence implies that whole spectrum methods to select features, such as PCA, may be dominated by information that is not chemical in origin. Hence, we have strongly pursued the metric approach in our work. There are two additional goals behind this process. First, relevant information is contained in a small number of points as compared to the total number of points in the entire absorption spectrum (dimensionality reduction). Second, information captured by metrics is spectroscopically understandable since they are selected by studying the biochemically significant features from the spectra.

The third step of the classification is to use this information to train a classifier by minimizing the errors or maximizing the detection probability of the classes. The selected features are used to predict the accuracy of classification using training data. The parameters of the algorithm, quantitative measures of accuracy as well as the feature set used to classify become apparent at this stage. Supervised recognition algorithms include several common ones based on Bayes rule, Artificial Neural Networks (ANN), Linear Discriminant Analysis (LDA) or Support Vector Machines (SVM). In general, the Bayes rule-based classifiers are most appropriate when a large training set in available while others typically estimate either the distribution of the data or the complexity of the contour required to separate classes in the dimensional space of the data. In our experience, the choice of a specific algorithm does not make a tremendous difference in the results obtained, however, the requirements for data size, speed and insight into classification differ substantially among the various algorithms. Finally, the validation set is employed to test the efficacy of the developed algorithm and present its results in an independent test. Sometimes, a part ($1/n$) of the training data may be designated as an independent validation set, the training repeated n times, each time leaving ($1/n$) part of the data for validation. This is termed n-fold cross-validation and is often used as a surrogate for a completely independent data set. We suggest that an independent data set be used and the n-fold technique be used as a preliminary test for histologic analyses due to the complexity of biological materials.

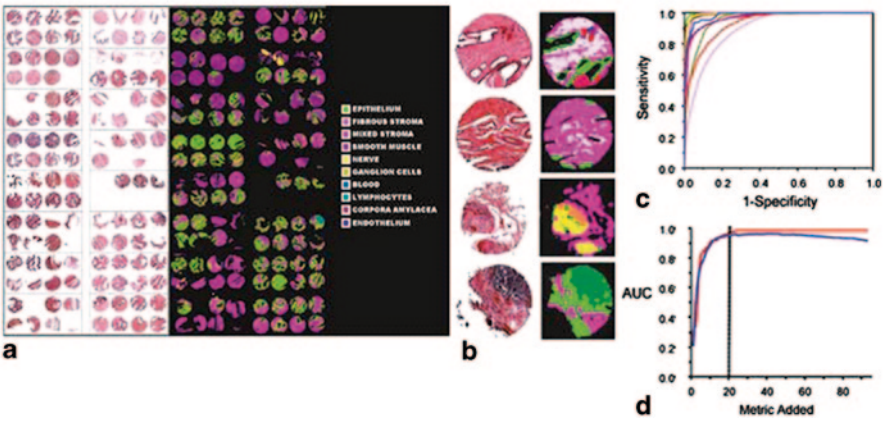


Fig. 16.9 **a** H&E stained images of prostate tissue are shown on the left and the classified infrared images are shown on the right. **b** Certain regions have been selected from the tissues, to show a correspondence between H&E and classified images. **c** ROC curves which depict the accuracy of classification. **d** The number of spectral features and classification accuracy has been shown in red for the training data and in blue for the validation data. (Reproduced from ref. [11] and ref. [48])

Let us examine an example of Supervised Learning in more detail. Figure 16.8 and will help in visualizing the steps involved. First we select a prostate tissue microarray as our data set. Next, the data are subjected to Minimum Noise Fraction for reducing noise, a scaled water vapor spectrum is employed to eliminate atmospheric interferences and spectra are all baseline corrected in a consistent manner. Absorbance images for some common features are employed to examine the data by comparing the various cell types of interest with a gold standard. Commonly, the gold standard is a visible microcopy image that is a corresponding tissue section stained with hematoxylin and eosin (H&E) and one in which a pathologist has recognized various entities of interest. Specialized immunohistochemical stains may also be employed, for example, smooth muscle-α (SMA) stain for identifying myofibroblasts, smooth muscles and myoepithelial tissue. The use of IHC may be difficult, however, and is not recommended to be used as the sole standard for a variety of reasons. Regions corresponding to different classes of cells are marked on the images as Regions of Interest (ROI) and the biochemical information inherent in ROI is extracted from the IR absorbance spectra. Potentially significant features from the absorption spectra are recognized, which defines the metrics to be used. Subsequently, computer algorithms are employed for training using the marked pixels as well as the selected metrics. Classification accuracy is optimized and statistics obtained. The developed classifier is finally run on the validation data set and quantitation of the accuracy of the classifier is obtained. We additionally test our data for robustness, for example, by decreasing signal to noise ratio until classifier fails, to determine the conditions under which it is bound. Finally, an optimized algorithm, metric set, calibration and validation statistics as well robustness analysis is obtained. Together, these tools can form the basis for automated histopathology (Fig. 16.9).

Unsupervised Learning Common unsupervised learning methods include K-Means and hierarchical clustering analysis. In this type of classification, prior knowledge of the classes within the data is not available and, indeed, the focus is often on discovering the underlying structure. The data acquired is divided by the algorithm used into clusters or groups. This form of classification is highly valuable as a visualization and discovery tool in IR imaging. The results from unsupervised methods may be used as is, especially when outliers are to be detected or a ready reference (such as a corresponding stained tissue section) is available. The routine use of these methods alone to extract biomedical information is usually not recommended as the visualization is strongly dependent on the specific sample(s) examined. Without a comprehensive analysis of many samples and deeper understanding of analysis at multiple levels of clustering, the effectiveness of these methods is significantly reduced. A secondary source of knowledge is required to validate the results of the classification, such as immunostaining or expert manual validation. When used as a discovery tool, results can be further translated with supervised methods to minimize errors and procure accurate results.

16.4 Application of Ft-IR in Histopathology, Especially in Cancer

One of the major areas of application for IR imaging has been in diagnosing, staging and grading cancer. A number of pioneering studies and reviews are available. For the sake of brevity, we focus here on salient features of a few studies. For convenience, these largely focus on work from our own laboratories and are only meant to be illustrative, not exhaustive or, necessarily, studies originating the ideas. We also do not focus on straightforward applications, i.e. applications in which two sets of tissues, one diseased and other normal, need to be separated. In our opinion, that is a rather simple application that can be a byproduct of more comprehensive protocols and is of limited practical value in itself. Hence, in our discussions of specific studies, we focus on interesting aspects of translation, studies in which exciting new information is available or where unexpected results were reported.

16.4.1 Breast Cancer

One of the most important applications of FT-IR spectroscopy in analyzing breast cancer tissues has been in trying to utilize FT-IR automated histopathology for classifying the cell types present in breast cancer tissues, in order to diagnose, grade and characterize the tumors. An example [50] of correlations between different cells and stains and IR imaging data is shown in Fig. 16.10. IR absorption spectra have been analyzed extensively for differentiating between the malignant and the benign breast tissues [51–53].

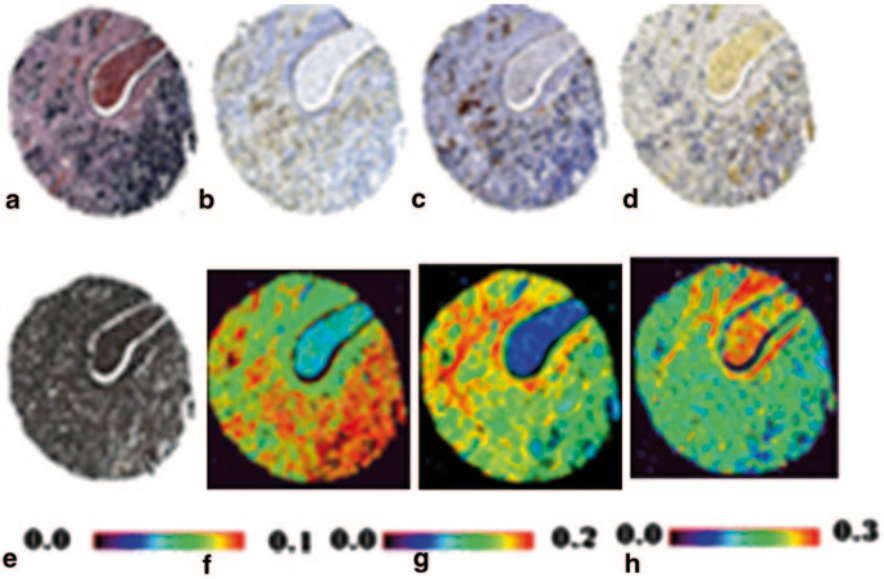


Fig. 16.10 Figure showing a comparison between H&E and IHC stains—clinical histopathology versus classified FT-IR images of breast tissue after automated histopathology: **a** H&E. **b** vimentin. **c** smooth muscle α actin (SMA). **d** factor VIII. **e** Unstained tissue. After automated histopathology—FT-IR image showing (**f**) highlighted epithelium at 1,080 cm^{-1}, **g** highlighted stroma at 1,236 cm^{-1} and **h** highlighted blood vessels at 3,294 cm^{-1}. (Reproduced from ref. [51])

ATR FT-IR imaging has been used recently to chemically characterize smaller-sized cells and finer morphologic structures such as endothelial cells, myoepithelial cells and terminal ductal lobular units, which are usually inaccessible by transmission mode imaging. In addition to smaller cells and structures, the approach was extended to study sub-cellular structures. In this study, chromosomes extracted from breast cancer cells were compared to normal cells for understanding the molecular changes introduced in DNA. FT-IR ATR spectra from the scalp and pubic hair samples from patients having breast cancer and people without cancer were also compared to differentiate between the molecular or chemical compositions in both cases; hence opening up yet another possibility of analyzing breast cancer [54].

IR spectroscopy has been applied to differentiate two cell lines—MCF-7 (estrogen receptor expressed, ER+) and SKBr3 (estrogen receptor non-expressed, ER-) both from the same breast cancer by the analysis of the chemical information available from the FT-IR absorption spectra [55]. An attempt has also been made by our group to characterize the stroma and study its transformations in epithelial tumor progression. One of the characteristics of cancer-associated stroma is the fibroblast-to-myofibroblast cellular transformation which is accompanied by the expression of α-smooth muscle actin (α-SMA). Transformation of fibroblast-to-myofibroblast was analysed in primary normal human dermal fibroblasts (NHDFs) activated with TGFβ1, in cocultures of primary fibroblasts with tumorigenic breast epithelial cells (MCF −7) and human tissues [56]. Moreover, in order to study the interaction be-

tween the cancerous epithelium and surrounding stroma for understanding early biomolecular events, which control the progression of the cancer, our group developed a 3D co-culture model from the MCF10A mammary acini model,along with a stroma simulated by primary human mammary fibroblasts, and co-culturing the stroma with MCF7 tumor breast epithelial cells [57]. FT-IR Spectroscopy has also been used to study the relationship between the chemical composition of microcalcifications in the breast tissue and the grade of breast cancer [58].

16.4.2 Prostate Cancer

As has been discussed in detail before, automated histopathology has been done on IR images of prostate tissues sections, to classify the different cell types—in order to detect as well as grade prostate cancer tissues [11, 59–61] and also to differentiate between 'confined' and 'locally invasive' prostate cancer [62]. R Spectra of prostate cancer tissue have been correlated with Gleason score and the grade of tumour as well [63]. FT-IR spectroscopy has been applied to study the effect of certain potent anticancer drugs on prostate cancer cells, in terms of chemical and metabolic changes to determine whether they would be effective, presenting an interesting adjunct set of data for future clinical protocols involving imaging [64–67]. Our group has also attempted to develop a completely automated histopathology technique for distinguishing between malignant and normal tissue. FT-IR spectroscopic images were recorded from an unstained tissue micro-array and then pattern recognition technique was used to identify epithelial cells without user input, by combining the IR and optical data. Morphological features, extracted by optimizing a minimum-redundancy-maximal-relevance (mRMR) criterion and sequential floating forward selection (SFFS), were applied to distinguish between malignant and normal tissues and showed almost perfect accuracy [59] (Fig. 16.11).

While the presence of disease can be determined, extensions of these concepts to detecting susceptibility to disease have also been carried out. For example, efforts have been made to extract chemical information from the absorption spectra of prostate adenocarcinoma tissues (CaP) from the two different populations–population with low susceptibility to CaP as India and population with high susceptibility to CaP as U.K., to recognize a susceptibility-to-adenocarcinoma biochemical signature [68] and differences between benign and cancerous tissue have also been demonstrated [69]. Post-diagnosis, the decision making in the clinic is often along the lines of determining risk based on grade and stage. Principal component-discriminant function analysis (PC-DFA) algorithm was applied to IR microspectroscopic images for CaP grading in which, using typical Gleason grading and the tumour/node/metastasis system, it was discovered, that the combined sensitivity and specificity of the three-band Gleason score criterion diagnosis was 92.3 and 99.4 % respectively. FT-IR spectroscopy along with histological stains have been used to analyse the biomolecular structures of metastatic CaP cells in the bone marrow, to understand the inherent biochemistry behind the proliferation of the cancerous cells at the metastatic sites, to help produce more effective anticancer treatments [70].

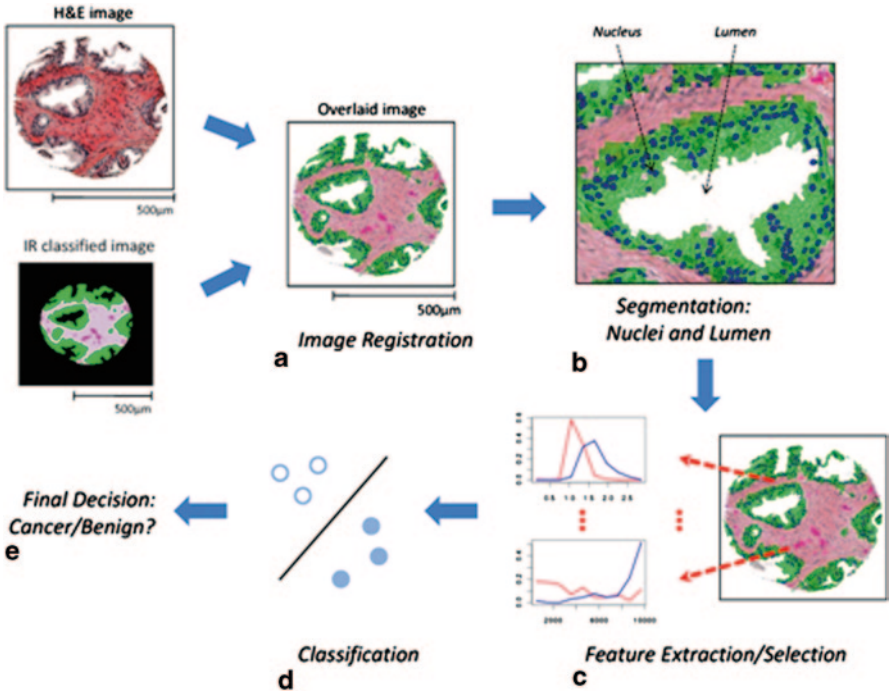


Fig. 16.11 **a** FT-IR classified image is overlaid with H&E stained image. **b** segmentation of nuclei and lumens in the tissue sample takes place next. **c** features are extracted and selected. **d** the tissue is classified. **e** diagnosis is done—whether the sample is cancerous or benign

16.4.3 Cancer of the Colon

FT-IR imaging has been used in several labs for diagnosing colon cancer by the chemometric analysis of the absorption spectra [71–73] distinguishing between healthy and malignant colon tissues and colon cancer and colitis [74]. Histopathological analysis of colon tissue samples has been done to classify the cell types in the colon tissues with high accuracy and specificity, in order to detect colon cancer as well as other diseases of the colon such as ulcerative colitis or Crohn's disease. Previously, transmission IR imaging along with Baysian classification had been used to classify the basic cell types like epithelial cells, goblet cells, loose connective tissue and blood cells [75]. Infrared images were then acquired employing the ATR FT-IR mode and Transmission high-definition (74×) mode, both of which showed much higher spatial resolution. The resulting images, as seen in, demonstrrate details that are not accessible at present. the intestinal crypts could be visualised much more clearly and even the small cells like lymphocytes and intracellular mucin (which are potential diagnostic markers for cancer) could be classified accurately. This new imaging capability is likely to prove most crucial for colon tissue. In prior applications, FT-IR absorption spectra and algorithms were employed in

distinguishing the colonic mucosa, between normal, hyperplasia, or dysplasia [76]. The biochemical modifications induced by CDCP1 (a transmembrane protein, believed to play a role in metastasis) expression in colon cancer tissue was also found to have a spectral signature [77] as did structural disorders in the DNA of colon cancer cells ([78]; Fig. 16.12).

16.4.4 Skin Cancer

Skin cancer is the most prevalent cancer and is quite amenable to analysis by IR imaging due to the characteristic structures present. IR spectroscopy has been applied to distinguish between benign, dysplastic and malignant skin cancers [79] and also between different types of non-melanoma skin cancers like basal cell carcinoma (BCC), squamous cell carcinoma (SCC) and Bowen's disease [80]. Melanoma diagnoses using common biomarkers separating them from normal skin tissues [81, 82] has also been accomplished. Two types of melanoma cell lines—murine melanoma (B16F10 cell line) and human melanoma (C8161 cell line) were also employed to provide important spectral correlations with disease [83]. In many cases, melanoma might originate from a nevus, which is actually a non-malignant chronic lesion on the skin. FTIR was also used to analyse and characterize the spectral differences between melanoma, nevus, and the normal epidermis [84]. A very interesting and recent development has been to use engineered tissues as models. When combined with spectroscopy and computation, a novel way for studying skin disease progression in eanbled. Our group has applied FT-IR imaging to examine an engineered tissue model of melanoma—the biochemical properties and spectral changes in different layers of growing skin were first characterized, then malignant melanocytes were introduced in the model, to simulate tumor formation and growth. Chemical changes associated with tumor-stromal interactions were then analysed during the course of tumor growth [85].

16.5 Conclusions

From the previous sections, it is abundantly evident that with the advancement of spectroscopic techniques and development of fast, robust computational algorithms, FT-IR spectroscopy is rapidly emerging in the modern world as a powerful diagnostic tool. It is being used in several biomedical and biochemical venues to derive information about the molecular structure of biological samples, identification of cell types and disease states. The ultimate aim is to adapt the technology to the clinic to help solve real life cases. A number of developments are underway towards that goal—high speed for data acquisition, effectively handling the large volume of data, high cost equipment, lower spatial resolution as compared to other microscopy approaches and lack of sufficiently robust analytic tools to deal with

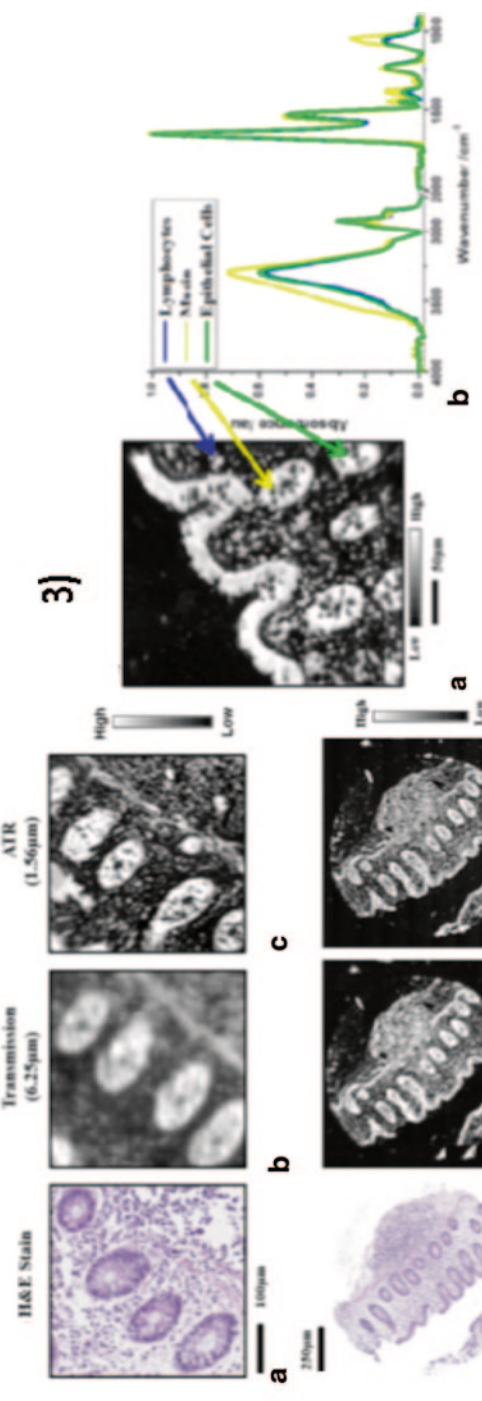


Fig. 16.12 **1a** H&E stain of a colon tissue section. **1b** Transmission-mode IR image (6.25 × 6.25 μm) of a similar section. **1c** ATR-mode IR image (1.56 × 1.56 μm) of a similar section. **2a** H&E stain of a colon tissue section. **2b** Transmission-mode image with a 15X, 0.5 NA objective. **2c** High-definition transmission mode image with a 74X, 0.65 NA objective. Both the ATR FT-IR as well as the HD mode image have much higher spectral resolution and the images are clearer than the Transmission mode images. **3a** epithelial cells, lymphocytes (*white dots*) and intracellular mucin (*black dots*) could be recognized in the ATR data. **3b** IR spectra corresponding to each of these celltypes. **1b**, **1c**, **2b**, **2c** and **3b**, **3c** all of them correspond to the CH2 asymmetric vibrational mode at 2,924 cm^{-1}. (Reproduced from ref. [76])

the highly complex and highly variable data sets—are all avenues in which remarkable progress is being demonstrated. Today's clinical pathology suffers from several limitations in the areas of diagnosis, prognosis as well as treatments of diseases. FT-IR histopathology has the potential of providing some advantages over common clinical histopathological techniques in all these areas of need.

References

1. Bhargava R (2012) Infrared spectroscopic imaging: the next generation. Appl Spectrosc 66(10):1091–120
2. Bhargava R (Oct (2007) Towards a practical Fourier transform infrared chemical imaging protocol for cancer histopathology. Anal Bioanal Chem 389(4):1155–1169
3. Srinivasan G, Bhargava R (Jul 2007) Fourier transform-infrared spectroscopic imaging: the emerging evolution from a microscopy tool to a cancer imaging modality. Spectroscopy 22(7):30
4. Levin IW, Bhargava R (5 May 2005) Fourier transform infrared vibrational spectroscopic imaging: integrating microscopy and molecular recognition. Annu Rev Phys Chem 56(1):429–474
5. Griffiths PR, de Haseth JA (2007) Fourier transform infrared spectrometry. Wiley, Hoboken
6. Jackson M (2004) Biomolecules to biodiagnostics: spectroscopy does it all. Faraday Discuss 126:1–18
7. Jackson M, Sowa MG, Mantsch HH (Oct 1997) Infrared spectroscopy: a new frontier in medicine. Biophys Chem 68(1–3):109–125
8. Diem M, Romeo M, Boydston-White S, Miljković M, Matthäus C (2004) A decade of vibrational micro-spectroscopy of human cells and tissue (1994-2004). Analyst 129(10):880–885
9. Martin FL, Kelly JG, Llabjani V, Martin-Hirsch PL, Patel II, Trevisan J, Fullwood NJ, Walsh MJ (2010) Distinguishing cell types or populations based on the computational analysis of their infrared spectra. Nat Prot 5:1748–1760
10. Diem M, Griffiths PR, Chalmers JM (eds) (2008) Vibrational spectroscopy for medical diagnosis. Wiley, Chichester
11. Fernandez DC, Bhargava R, Hewitt SM, Levin IW (2005) Infrared spectroscopic imaging for histopathologic recognition. Nat Biotechnol 23(4):469–474
12. Colarusso P, Kidder LH, Levin IW, Fraser JC, Arens JF, Lewis EN (1998) Infrared spectroscopic imaging: from planetary to cellular systems. Appl Spectrosc 52:106a–120a
13. Walsh MJ, Reddy RK, Bhargava R (2012) Label-free biomedical imaging with mid-Infrared microspectroscopy. IEEE J Sel Top Quant 18:1502–1513
14. Dumas P, Jamin N, Teillaud JL, Miller LM, Beccard B (2004) Imaging capabilities of synchrotron infrared microspectroscopy. Faraday Discuss 126:289–302. discussion 303–311
15. Miller LM, Dumas P (Jul 2006) Chemical imaging of biological tissue with synchrotron infrared light. Bio chim Bio phys Acta 1758(7):846–857
16. Dumas P, Sockalingum GD, Sule-Suso J (Jan 2007) Adding synchrotron radiation to infrared microspectroscopy: what's new in biomedical applications? Trends Biotechnol 25(1):40–44
17. Hirschmugl CJ, Gough KM (May 2012) Fourier transform infrared spectrochemical imaging: review of design and applications with a focal plane array and multiple beam synchrotron radiation source. Appl Spectrosc 66(5):475–491
18. Bhargava R, Levin IW (eds) (2005) Spectrochemical analysis using infrared multichannel detectors. Blackwell Publishing, Oxford, pp 56–84
19. Lewis EN, Treado PJ, Reeder RC, Story GM, Dowrey AE, Marcott C, Levin IW (1995) Fourier transform spectroscopic imaging using an infrared focal-plane array detector. Anal Chem 67:3377–3381

20. Bhargava R, Wall BG, Koenig JL (2000) Comparison of the FT-IR mapping and imaging techniques applied to polymeric systems. Appl Spectrosc 54:470–474
21. Koenig JL, Wang SQ, Bhargava R (2001) FT-IR Images. Anal Chem 73:360A–369A
22. Bhargava R, Levin IW (2003) Time-resolved Fourier transform infrared spectroscopic imaging. Appl Spectrosc 57:357–366
23. Bhargava R, Levin IW (2003) Noninvasive imaging of molecular dynamics in heterogeneous materials. Macromolecules 36:92–96
24. Bhargava R, Levin IW (2001) Fourier transform infrared imaging: theory and practice. Anal Chem 73:5157–5167
25. Nasse MJ, Walsh MJ, Mattson EC, Reininger R, Kajdacsy-Balla A, Macias V et al (May 2011) High-resolution fourier-transform infrared chemical imaging with multiple synchrotron beams. Nat Methods 8(5):413–416
26. Walsh MJ, Mayerich D ,Kajdacsy-Balla A, Bhargava R (2012) High resolution mid-infrared Imaging for disease diagnosis. Proc. of SPIE Vol. 8219 82190R–1
27. Reddy RK, Walsh MJ, Schulmerich MV, Carney PS, Bhargava R (2013) High-definition infrared spectroscopic imaging. Appl Spectrosc 67:93–105
28. Nasse MJ, Mattson E, Hirschmugl C (2010) First results from IRENI—rapid diffraction-limited high resolution imaging across the mid-infrared bandwidth. AIP Conference Proceedings 1214, pp 108–110
29. Stavitski E, Smith RJ, Bourassa MW, Acerbo AS, Carr GL, Miller LM (2013) Dynamic full-field infrared imaging with multiple synchrotron beams. Anal Chem 85(7):3599–3605
30. Marcelli A, Cricenti A, Kwiatek WM, Petibois C (2012) Biological applications of synchrotron radiation infrared spectromicroscopy. Biotechnology Advances 30(6):1390–1404
31. Petibois C, Cestelli-Guidi M, Piccinini M, Moenner M, Marcelli A (2010) Synchrotron radiation FTIR imaging in minutes: a first step towards real-time cell imaging. Anal Bioanal Chem 397(6):2123–2129
32. Marcelli A, Cinque G (2011) Infrared synchrotron radiation beamlines: high brilliance tools for IR spectromicroscopy. RSC Analytical Spectroscopy Series, pp 67–104
33. Walsh MJ, Holton SE, Kajdacsy-Balla A, Bhargava R (2012) Attenuated total reflectance fourier transform infrared spectroscopic imaging for comprehensive breast tissue histopathology. Vib Spectrosc 60:23–28
34. Kazarian SG, Chan KLA (2006) Applications of ATR-FTIR spectroscopic imaging to biomedical samples. Biochim Biophys Acta—Biomembranes 1758(7):858–867
35. Kazarian SG, Chan KLA (2013) ATR-FTIR spectroscopic imaging: recent advances and applications to biological systems. Analyst 138(7):1940–1951
36. Bassan P, Byrne HJ, Bonnier F, Lee J, Dumas P, Gardner P (2009) Resonant Mie scattering in infrared spectroscopy of biological materials—understanding the 'dispersion artefact'. Analyst 134(8):1586–1593
37. Lee J, Gazi E, Dwyer J, Brown MD, Clarke NW, Nicholson JM, Gardner P (2007) Optical artefacts in transflection mode FTIR microspectroscopic images of single cells on a biological support: the effect of back-scattering into collection optics. Analyst 132(8):750–755
38. Bassan P, Byrne HJ, Lee J, Bonnier F, Clarke C, Dumas P, Gazi E et al (2009) Reflection contributions to the dispersion artefact in FTIR spectra of single biological cells. Analyst 134(6):1171–1175
39. Romeo M, Mohlenhoff B, Diem M (2006) Infrared micro-spectroscopy of human cells: causes for the spectral variance of oral mucosa (buccal) cells. Vib Spectrosc 42(1):9–14
40. Filik J, Frogley MD, Pijanka JK, Wehbe K, Cinque G (2012) Electric field standing wave artefacts in FTIR micro-spectroscopy of biological materials. Analyst 137(4):853–861
41. Wehbe K, Filik J, Frogley MD, Cinque G (2013) The effect of optical substrates on micro-FTIR analysis of single mammalian cells. Anal Bioanal Chem 405(4):1311–1324
42. Gulley-Stahl HJ, Bledsoe SB, Evan AP, Sommer AJ (2010) The advantages of an attenuated total internal reflection infrared microspectroscopic imaging approach for kidney biopsy analysis. Appl Spectrosc 64(1):15–22

43. Davis BJ, Carney PS, Bhargava R (2010) Theory of mid-infrared absorption microspectroscopy. II. Heterogeneous samples. Anal Chem 82:3487–3499
44. Davis BJ, Carney PS, Bhargava R (2010) Theory of mid-infrared absorption microspectroscopy. I. Homogeneous samples. Anal Chem 82:3474–3486
45. Davis BJ, Carney PS, Bhargava R (2011) Infrared microspectroscopy of intact fibers. Anal Chem 83:525–532
46. van Dijk T, Mayerich D, Carney PS, Bhargava R (2013) Recovery of absorption spectra from Fourier transform infrared microspectroscopic measurements of intact spheres. Appl Spectrosc 67:546–552
47. Miller LM, Dumas P (Jul 2006) Chemical imaging of biological tissue with synchrotron infrared light. Biochim Biophys Acta (BBA)—Biomembranes 1758(7):846-857, ISSN 0005-2736, http://dx.doi.org/10.1016/j.bbamem.2006.04.010
48. Bhargava R, Fernandez DC, Hewitt SM, Levin IW (2006) High throughput assessment of cells and tissues: Bayesian classification of spectral metrics from infrared vibrational spectroscopic imaging data. Biochim Biophys Acta 1758:830–845
49. Kwak JT, Reddy RK, Sinha S, Bhargava R (2012) An analysis of the sources of variance in Fourier transform infrared spectroscopic imaging of tissues. Anal Chem 84:1063–1069
50. Pounder NF et al. Proc. of SPIE vol 7182 718206–2
51. Fabian H, Lasch P, Boese M, Haensch W (16 Dec 2003) Infrared microspectroscopic imaging of benign breast tumor tissue sections. J Mol Struct 661:411–417
52. Fabian H, Lasch P, Boese M, Haensch W (2002) Mid-IR microspectroscopic imaging of breast tumor tissue sections. Biopolymers 67(4–5):354–357
53. Fabian H, Thi NAN, Eiden M, Lasch P, Schmitt J, Naumann D (Jul 2006) Diagnosing benign and malignant lesions in breast tissue sections by using IR-microspectroscopy. Biochimica Et Biophysica Acta-Biomembranes 1758(7):874–882
54. Lyman DJ, Murray-Wijelath J (Jan 2005) Fourier transform infrared attenuated total reflection analysis of human hair: comparison of hair from breast cancer patients with hair from healthy subjects. Appl Spectrosc 59(1):26–32
55. BüttnerMostaço-Guidolin L, Murakami LS, RibeiroBatistuti M, Nomizo A, Bachmann L (2010) Molecular and chemical characterization by Fourier transform infrared spectroscopy of human breast cancer cells with estrogen receptor expressed and not expressed. Spectroscopy 24:501–510
56. Holton SE, Walsh MJ, Kajdacsy-Balla A, Bhargava R (21 Sept 2011) Label-free characterization of cancer-activated fibroblasts using infrared spectroscopic imaging. Biophys J 101(6):1513–1521
57. Holton SE, Bergamaschi A, Katzenellenbogen BS, Bhargava R (2012) A spectroscopic signature associated with hormone sensitivity in 3D co-culture models of breast cancer. AACR 103rd annual meeting, Chicago, March 31-April 4 2012
58. Baker R, Rogers KD, Shepherd N, Stone N (28 Sept 2010) New relationships between breast microcalcifications and cancer. Br J Cancer 103(7):1034–1039
59. Kwak JT, Hewitt SM, Sinha S, Bhargava R (9 Feb 2011) Multimodal microscopy for automated histologic analysis of prostate cancer. BMC Cancer 11:62
60. Baker MJ, Gazi E, Brown MD, Shanks JH, Clarke NW, Gardner P (Feb 2009) Investigating FTIR based histopathology for the diagnosis of prostate cancer. J Biophotonics 2(1–2):104–113
61. Mackanos MA, Contag CH (Dec 2009) FTIR microspectroscopy for improved prostate cancer diagnosis. Trends Biotechnol 27(12):661–663
62. Baker MJ, Gazi E, Brown MD, Shanks JH, Gardner P, Clarke NW (25 Nov 2008) FTIR-based spectroscopic analysis in the identification of clinically aggressive prostate cancer. Br J Cancer 99(11):1859–1866
63. Gazi E, Baker M, Dwyer J, Lockyer NP, Gardner P, Shanks JH et al (Oct 2006) A correlation of FTIR spectra derived from prostate cancer biopsies with gleason grade and tumour stage. Eur Urol 50(4):750–761
64. Gasper R, Goormaghtigh E (Dec 2010) Effects of the confluence rate on the FTIR spectrum of PC-3 prostate cancer cells in culture. Analyst 135(12):3048–3051

65. Gasper R, Mijatovic T, Benard A, Derenne A, Kiss R, Goormaghtigh E (Nov 2010) FTIR spectral signature of the effect of cardiotonic steroids with antitumoral properties on a prostate cancer cell line. Biochim Biophys Acta 1802(11):1087–1094
66. Gasper R, Mijatovic T, Kiss R, Goormaghtigh E (1 Jan 2010) FTIR spectroscopy reveals the concentration dependence of cellular modifications induced by anticancer drugs. Spectrosc Int J 24(1):45–49
67. Derenne A, Gasper R, Goormaghtigh E (21 Mar 2011) The FTIR spectrum of prostate cancer cells allows the classification of anticancer drugs according to their mode of action. Analyst 136(6):1134–1141
68. Patel II, Trevisan J, Singh PB, Nicholson CM, Krishnan RK, Matanhelia SS et al (Aug 2011) Segregation of human prostate tissues classified high-risk (UK) versus low-risk (india) for adenocarcinoma using fourier-transform infrared or ramanmicrospectroscopy coupled with discriminant analysis. Anal Bioanal Chem 401(3):969–982
69. Gazi E, Dwyer J, Gardner P, Ghanbari-Siahkali A, Wade AP, Miyan J et al (Sep 2003) Applications of fourier transform infrared microspectroscopy in studies of benign prostate and prostate cancer. A pilot study. J Pathol 201(1):99–108
70. Gazi E, Dwyer J, Lockyer NP, Gardner P, Shanks JH, Roulson J et al (Mar 2007) Biomolecular profiling of metastatic prostate cancer cells in bone marrow tissue using FTIR microspectroscopy: a pilot study. Anal Bioanal Chem 387(5):1621–1623
71. Khanmohammadi M, Garmarudi AB, Ghasemi K, Jaliseh HK, Kaviani A (2009) Diagnosis of colon cancer by attenuated total reflectance-fourier transform infrared microspectroscopy and soft independent modeling of class analogy. Med Oncol 26(3):292–297
72. Khanmohammadi M, BagheriGarmarudi A, Samani S, Ghasemi K, Ashuri A (Jun 2011) Application of linear discriminant analysis and attenuated total reflectance fourier transform infrared microspectroscopy for diagnosis of colon cancer. Pathol Oncol Res 17(2):435–441
73. Lasch P, Haensch W, Lewis EN, Kidder LH, Naumann D (Jan 2002) Characterization of colorectal adenocarcinoma sections by spatially resolved FT-IR microspectroscopy. Appl Spectrosc 56(1):1–9
74. Li X, Li QB, Zhang GJ, Xu YZ, Sun XJ, Shi JS et al (2012) Identification of colitis and cancer in colon biopsies by fourier transform infrared spectroscopy and chemometrics. Sci World J 2012:936149
75. Walsh MJ et al Proc of SPIE Vol 8219 82190R–2
76. Mackanos MA, Hargrove J, Wolters R, Du CB, Friedland S, Soetikno RM et al (Jul-Aug 2009) Use of an endoscope-compatible probe to detect colonic dysplasia with fourier transform infrared spectroscopy. J Biomed Opt 14(4):044006
77. Yousef I, Breard J, SidAhmed-Adrar N, Maamer-Azzabi A, Marchal C, Dumas P et al (21 Dec 2011) Infrared spectral signatures of CDCP1-induced effects in colon carcinoma cells. Analyst 136(24):5162–5168
78. Kondepati VR, Heise HM, Oszinda T, Mueller R, Keese M, Backhaus J (11 Mar 2008) Detection of structural disorders in colorectal cancer DNA with fourier-transform infrared spectroscopy. Vib Spectrosc 46(2):150–157
79. Tosi G, Conti C, Giorgini E, Ferraris P, Garavaglia MG, Sabbatini S et al FTIR microspectroscopy of melanocytic skin lesions: a preliminary study. Analyst (12):3213
80. Ly E, Piot O, Durlach A, Bernard P, Manfait M (Jun 2009) Differential diagnosis of cutaneous carcinomas by infrared spectral micro-imaging combined with pattern recognition. Analyst 134(6):1208–1214
81. Hammody Z, Argov S, Sahu RK, Cagnano E, Moreh R, Mordechai S (Mar 2008) Distinction of malignant melanoma and epidermis using IR micro-spectroscopy and statistical methods. Analyst 133(3):372–378
82. Mordechai S, Sahu RK, Hammody Z, Mark S, Kantarovich K, Guterman H et al (Jul 2004) Possible common biomarkers from FTIR microspectroscopy of cervical cancer and melanoma. J Microsc Oxf 215:86–91

83. Mostaço-Guidolin LB, Murakami LS, Nomizo A, Bachmann L (23 Jul 2009, Nov 2012) Fourier transform infrared spectroscopy of skin cancer cells and tissues. Appl Spectrosc Rev 44(5):438–455
84. Hammody Z, Sahu RK, Mordechai S, Cagnano E, Argov S (18 Mar 2005) Characterization of malignant melanoma using vibrational spectroscopy. Sci World J 5:173–182
85. Kong R, Reddy RK, Bhargava R (Jul 2010) Characterization of tumor progression in engineered tissue using infrared spectroscopic imaging. Analyst 135(7):1569–1578

Chapter 17
Vibrational Microspectroscopy for Analysis of Atherosclerotic Arteries

T.P. Wróbel, A. Fedorowicz, Ł. Mateuszuk, E. Maślak, A. Jasztal, S. Chlopicki and K.M. Marzec

Abstract This Chapter describes the fundamental principles of atherosclerotic plaque formation along with features detectable by histology. The main focus is on reviewing the use of Raman and IR spectroscopy in the study of atherosclerotic arteries. The analysis of the experimental results is possible only with the application of mathematical and statistical tools of analysis, i.e. chemometrics, which are also described. This chapter proves that chemometrics allows us to create a link between IR and Raman spectroscopy and their application in chemistry.

Keywords Infrared spectroscopy • Raman spectroscopy • Chemical imaging • Chemometrics • Atherosclerosis • Atherosclerotic plaque • Cardiovascular disease

17.1 Introduction

The desire of scientists to have an insight into the biochemical composition of a structure of tissues without the need for labelling, has initiated a number of studies utilizing the application of various spectroscopic techniques [1]. Raman and IR spectroscopies were found to be among the most successful for studying the differences in the relative concentrations of specific groups of compounds. They have been used successfully in many studies where the goal was to differentiate between normal and pathological tissue [2]. The use of Fourier Transform Infrared Spectroscopy (FT-IR) in biomedical sciences has increased significantly in

K.M. Marzec (✉) T.P. Wróbel · A. Fedorowicz · Ł. Mateuszuk · E. Maślak · A. Jasztal · S. Chlopicki
Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Bobrzynskiego 14, 30-348, Krakow, Poland
e-mail: marzec@chemia.uj.edu.pl

T.P. Wróbel
Faculty of Chemistry, Jagiellonian University, Ingardena 3, 30-060, Krakow, Poland

S. Chlopicki
Department of Experimental Pharmacology, Chair of Pharmacology, Jagiellonian University, Medical College, Grzegorzecka 16, 31-531, Krakow, Poland

M. Baranska (ed.), *Optical Spectroscopy and Computational Methods in Biology and Medicine,* Challenges and Advances in Computational Chemistry and Physics 14,
DOI 10.1007/978-94-007-7832-0_17, © Springer Science+Business Media Dordrecht 2014

recent years [3–7].The biggest field of FT-IR imaging applications relate to the analysis of tumours and their progression [4, 8–10] and to the field of cellular contents analysis [11–14]. All of these applications are based on the fact that FT-IR spectroscopy is a very sensitive technique, which when further supported by a proper data analysis approach, reveals great potential, the full scope of which is still being developed. In addition to this, recent advancements allow for the investigation of the spatial content distribution of samples, giving scientists new and powerful tools of research. Studies with the use of Raman spectroscopy were carried out *in vitro* in order to study biochemical changes observed during urological pathologies [15], gastric and lung cancers [16, 17], as well as changes in metabolite components in blood [18]. There are also studies in which Raman spectroscopy has been used *in vivo*. Some examples include *in vivo* detection of cervical precancers [19], *in vivo* studies where the feasibility of near-infrared Raman spectroscopy was presented during clinical gastrointestinal endoscopy [20], *in vivo* imaging of heart [21] and skin [22]. Other important clinical applications of vibrational spectroscopy include measurements of blood analytes that might prove useful in diagnosis of such diseases as breast cancer, Alzheimer's disease and atherosclerosis based on specific pattern of blood spectroscopic signals [23]. One of the important applications of these techniques that have emerged recently relates to *in vitro* and *in vivo* imaging that allows for qualitative and quantitative analysis of atherosclerotic plaques.

17.2 Atherosclerotic Arteries

17.2.1 Aorta—Biochemistry and Anatomy

Arteries are composed of three layers: *tunica adventitia*, *tunica media* and *tunica intima* (Fig. 17.1a).

The most typical components of *tunica adventitia,* that is the outer layer of the vessel wall, are type I collagen—networks of elastic fibers and proteoglycans that form loose connective tissue [24, 25]. Within adventitia there are scattered fibroblasts, adipocytes, rarely macrophages, lymphocytes and mast cells. The adventitia also contains unmyelinated nerve fibers and, in large arteries, small vessels called *vasa vasorum* [26].

The middle arterial layer, *tunica media*, consists of diagonally arranged smooth muscle cell sublayers separated by wavy sheets of elastic membranes [24, 25]. The major components of this layer are elastin, fibers of collagen type III, proteoglycans and glycoproteins (secreted by vascular smooth-muscle cells) [25, 26].

The innermost layer, *tunica intima*, is made up of a single-cell layer of endothelial cells, subendothelial space and internal elastic lamina made of elastin, that constitutes the border between intima and media. In the subendothelial space, there are fibrils of elastin and extracellular matrix [25, 26].

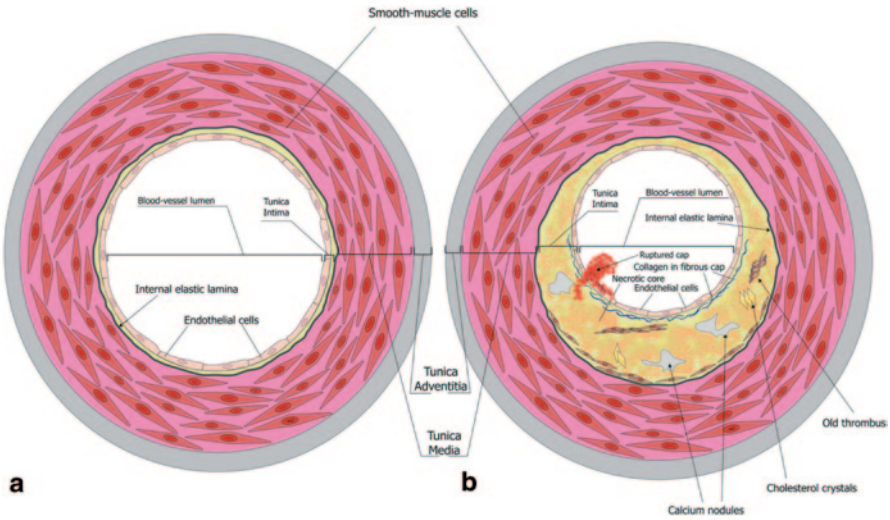


Fig. 17.1 Schematic representation of vascular wall. **a** Normal artery, **b** Artery with the most important features of advanced atherosclerotic plaque that could be detected with the use of vibrational spectroscopy

17.2.2 Atherosclerosis

Atherosclerosis (Atherothrombosis) is a major cause of ischaemic heart disease, stroke and peripheral vascular disease and, therefore, a leading cause of death. Although it was believed that atherosclerotic plaque resulted from accumulation of excessive plasma lipoproteins, progressively reducing the lumen of vessels it is now clear that it is as a chronic inflammatory disease of the vascular wall that is triggered and promoted by the development of endothelial dysfunction. These mechanisms are far from being entirely understood. However, it is becoming increasingly clear that Low-density lipoproteins (LDL) and their subsequent enzymatic or non-enzymatic oxidation to oxidised LDL (ox-LDL) and accumulation in subendothelial space triggers vascular inflammatory response the hallmark of which is endothelial activation and dysfunction [27, 28]. The development of endothelial dysfunction results in infiltration of immune inflammatory cells (monocytes, lymphocytes) to subendothelial space. Monocytes differentiate into macrophages (and/or dendritic cells) which phagocyte oxLDL and in consequence lose their mobility and converts to foam cells. Both foam cells and T-lymphocytes are detected in fatty streaks— asymptomatic, easy reversible early predecessor of atherosclerotic plaque. Formation of fatty streaks occurs in predilection sites with disturbed laminar flow (i.e. arteries' bifurcations) [29–34]. Atherosclerotic plaque formation involves extensive production of extracellular matrix. Both foam cells, activated endothelium and vascular smooth muscle cells produce increased amount of extracellular matrix in vascular wall which binds LDL subsequently beeing modified to oxLDL.

[27, 28]. Immobilized foam cells, lipids, debris, crystals of cholesterol (lipid core) surrounded by vascular smooth muscle cells, connective tissue T-lymphocytes and other elements accumulated in subendothelial space form assymetrical focal thickening of *tunica media*—atherosclerotic plaque [29–31, 35].

Further activation of inflammatory processes in atherosclerotic plaque results in matrix metalloproteinases (MMPs), cysteine proteases and mast cell- derived chymase secretion. This may lead to thinning of cap of atherosclerotic plaque, which is prone to rupture [36–39] and form so called vulnerable atherosclerotic plaque. Thin cap surrounding the whole lesion structure contains a smaller amount of smooth muscle cells, extracellular matrix components and is highly infiltrated by activated macrophages Fig. 17.1b. A plaque rupture can be defined as a disruption of the fibrous cap accompanied by intrusion of erythrocytes into the plaque itself [40]. Plague rupture leads to intravascular thrombus formation and partial or total artery occlusion. Plaque ruptures do not always result in ischemic symptoms and are often clinically silent. In advanced plaque angiogenesis is actived but new vessels are susceptible for rupture. Intraplaque haemorrhage occurring during plaque rupture may be detected in healed lesion by the presence of erythrocytes surrounded by lipid core components [41]. The majority of unstable plaque structure is composed of large "necrotic core", occupying approximately 1/3 to ½ of the total plaque area, whereas in unruptured vulnerable plaques it occupies usually less than ¼ of the lesion [42]. Typical for this structure is large free cholesterol content (including cholesterol crystals) and spotty calcifications, whereas in contrast to early plaques cholesterol esters compartment is significantly smaller. Necrotic core is a major part of plaque in advanced atherosclerosis. Massive necrosis within atherosclerotic plaque leads to the accumulation of cellular remnants in necrotic core area, which becomes "the graveyard of dead macrophages" [43] and the major pro-inflammatory stimuli that promotes vascular inflammation. The model of plaque's features important in atherosclerotic plaque progression (described above) and detectable with the use of vibrational spectroscopies is presented in Fig. 17.1b.

17.2.3 Techniques in the Diagnosis of Atherosclerosis

There are numerous amounts of imaging techniques in vascular biology and medicine that could by applied for atherosclerotic plaque imaging. A list of the most common imaging techniques includes: integrated backscatter intravascular ultrasound (IB-IVUS), virtual histology intravascular ultrasound (VH-IVUS), optical coherence topography (OCT), magnetic resonance imaging (MRI), angiography, thermography, angioscopy, elastography, positron emission topography (PET) and electron-beam computed topography (EBCT).

One of the most interesting trends is the use of laser spectroscopy for identification of atherosclerotic plaques. An example of this is the use of fluorescence spectroscopy (IFS) in the identification and diagnosis of atherosclerotic lesions [44, 45] and guidance for laser angioplasty systems [46]. Other optical diagnostics techniques, which can be used are reflectance spectroscopy (DRS) [45] and light scat-

tering and optical coherent topography [23, 47]. At present, techniques used in the clinical diagnosis of atherosclerosis are not able to detect vulnerable atherosclerotic plaques [48, 49], however there are new imaging methods which could be successfully use to detect high-risk plaque prone for rapture. These include intravascular ultrasound, intravascular magnetic resonance imaging (IMRI), thermography, angioscopy, elastography, positron emission topography and electron-beam computed topography [50].The imaging techniques (eg. CT and MRI), noninvasive electrophysiological tests (for vulnerable myocardium) and emerging catheters, as well as genomic and proteomic techniques, have also been presented as methods which could guide us in the search for vulnerable plaque [51].

The measurement of atherosclerotic aorta tissue with the use of vibrational spectroscopy dates back to 1985 [52], yet the results reported there are of uncertain quality due to the method of sample preparation. The first three significant papers on atherosclerosis imaging by FT-IR spectroscopy were published in 1991 by a Baraga et al. [53] and by Kodali et al. [54] and just two years later in 1993 by the former group [55]. Then in 1997 Hayashi et al. [56] analyzed rabbit vascular tissue after photodynamic therapy. Following this, in 2002 Li et al. [57] published biochemical analysis of human aorta containing atherosclerotic plaques. Later, in 2003 Adar et al. [58] reported on atherosclerosis progression in apoE mice with the use of vibrational spectroscopy. A year later a multi-technique based characterization of atherosclerotic tissue was performed by Becker et al. [59] Simultaneously, Wetzel et al. [60] studied collagen types in human atheroma with synchrotron radiation. At that time the imaging equipment was sufficiently developed to allow the first high quality imaging to be done in 2004 by Colley et al., [61] which was followed a year later by a review done by the same group led by prof. Kazarian [5].

In the following year, Wang et al. [62] reported the possibility of using FT-IR for tissue classification purposes *ex vivo*, and two years later (2008), the same group published a review also covering atherosclerotic tissue FT-IR measurements [4]. 2009 brought three articles by Palombo et al. in which vibrational spectroscopy was used to characterize the effects of iNOS deletion [63] and L-arginine treatment [64] on atherosclerosis progression. Palombo et al. studied also drug diffusion into the aortic wall using this technique [65]. Independently, Ebenstein et al. [66] correlated spectral features with mechanical properties of atherosclerotic tissue. In 2011, Wrobel et al. [67], released work focused on lipids distribution analysis in atherosclerotic plaque, and a year later work about protein profile changes in atherosclerotic plaque [68]. Finally, in 2012, Latermann et al. [69] investigated atherosclerotic plaques of rabbits with a complementary approach of FT-IR and Raman spectroscopies. Other works featured the differentiating of the atherosclerotic plaques in arterial tissues was done with the use of the multimodal coherent anti-Stokes Raman Scattering (CARS) microscopy [70–72].

We could take a risk and say that the future belongs to multimodal spectroscopy (MMS). As an example of this, we reported on can present the combination of DRS, IFS and Raman spectroscopy, [73] which allow for depth sensing. DRS provides information about the absorption and scattering properties of tissues and gives information mainly about hemoglobin and ceroid fraction, while IFS provides information about tissue fluorophores like collagen, elastin and lipids [73].

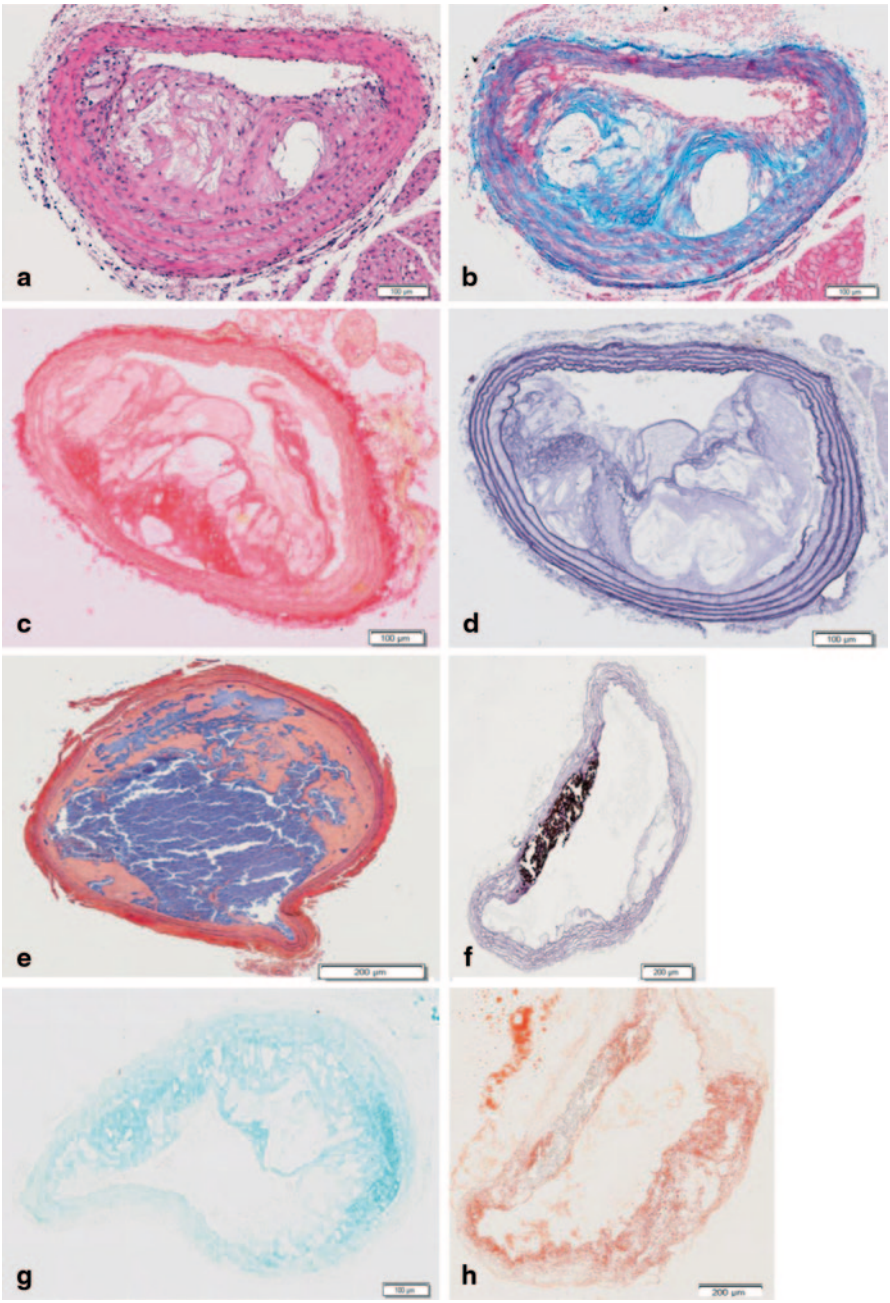


Fig. 17.2 **a–d**, **f–h** Brachiocephalic artery of ApoE/LDLR$^{-/-}$ mice. **a** Hematoxylin and eosin staining method for general morphology. **b** Collagen fibers distribution (*blue*), erythrocytes (*deep red*) and myofibroblasts (*pink/red*) localization stained with Gomori's one step trichrome method. **c** Collagen fibers (*red*) and muscles (*yellow*) stained with Picro Sirius Red method. **d** Elastic fibers

Both IR and Raman spectroscopic images are label-free microscopy techniques. However, they should be compared with histological staining. Such a comparison allows for the better understanding of spectroscopic results. Therefore, the following chapter discusses the most common histological stainings used in the assessment of atherosclerotic plaque.

17.2.4 Histological Examination and Immunostaining of Atherosclerotic Plaque

Atherosclerotic plaque is a heterogeneous structure. The differences occur in the localization and distribution of the cells and extracellular matrix. The intraplaque heterogeneity requires the use of special preparation and histological methods to characterize the morphology and composition of the plaque. The paraffin method is the most used in tissue preparation, however to visualize fats and other lipids frozen sections are necessary. The most common staining techniques with the examples are presented in Fig. 17.2.

The hematoxylin and eosin (H&E) staining method is useful to visualize the morphology and cellular heterogeneity of the plaque. Hallmarks of the atherosclerotic plaque that can be specified by H&E staining include the presence of macrophages and foam cells as well as the extent of the inflammatory and red blood cells (Fig. 17.2a). Very useful staining method for visualization of plaque morphology is also Gomori's One Step Trichrome. The trichrome method additionally shows collagen fiber deposits and allows for the determination of myofibroblasts and erythrocytes (Fig. 17.2b). Another staining method, which is dedicated to establishing collagen fiber distribution, is Picro Sirius Red. This procedure can also be used for the characterization of muscle cells (Fig. 17.2c). The Resorcin-Fuchsin staining method is used to visualize different fibrous proteins of extracellular matrix of elastin in atherosclerotic plaque (Fig. 17.2d). By using the PTAH (phosphotungstic acid hematoxylin) method, erythrocytes and fibrin deposition can be detected (Fig. 17.2e). Accordingly, with the progression of atherosclerotic plaque, calcification and matrix proteoglycans may occur. To detect calcium in atherosclerotic plaques, the Von Kossa staining is considered to be one of the most commonly used methods (Fig. 17.2f). The Alcian Blue method of staining is recommended for detection of matrix proteoglycans (Fig. 17.2). Lipids are the components that require frozen tissue preparation. They appear in the foam cells and in the extracellular matrix forming the lipid core. The Oil Red O staining method is routinely used to establish its presence in the plaque (Fig. 17.2h) [74–75].

(*violet*) stained with resorcin-fuchsin method. **f** Calcium deposition (*black*) stained with Van Koss method. **g** Matrix proteoglycans (*intensive blue*) stained with alcian blue method. **h** Lipids (*red*) stained with Oil Red O method. **e** Femoral artery of Wistar rat. Fibrin deposition (*light blue*) and erythrocytes (*dark blue*) stained with PTAH method

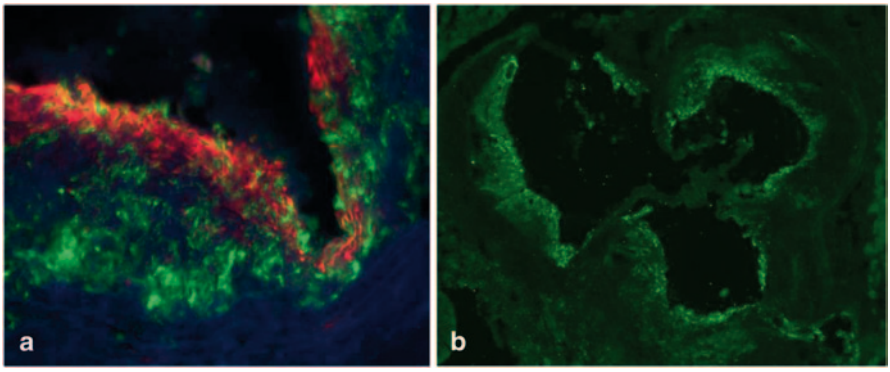


Fig. 17.3 **a** Double immunostaining of atherosclerotic plaque. Slide was acetone fixed, then Cy3-conjugated anti α- actin Ig (orange fluorescence) and rabbit anti CD68 Ig were used. At the second step, Alexa Fluor 488- conjugated goat anti rabbit Ig (green fluorescence) was used to visualize CD68- positive macrophages in plaque area. The fibrous cap is highly infiltrated by macrophages, which may suggest the tendency to plaque rupture. **b** Cross section of immunostained aortic root isolated from atherosclerotic mouse (C57Bl/6J ApoE/LDLr$^{-/-}$). The area of three advanced atherosclerotic plaques is mostly composed of CD68- positive macrophages and macrophage remnants, visualized by FITC- conjugated antibody (green fluorescence)

Atherosclerotic plaque components can be also visualized by immunohistochemical (IHC) staining (Fig. 17.3). IHC staining localizes the antigen in tissue sections with specific antibody (through antigen-antibody interactions) labeled with a marker such as fluorescent dye (e.g. Cy3, FITC or DAPI), enzyme (e.g. alkaline phosphatase or horseradish peroxidase) or radioactive particle. The tissue containing atherosclerotic plaques requires fixation for optimal specimen preservation. The most common is paraformaldehyde-based fixation but more safer procedure for antigen preservation is fixing in acetone followed with stiffening of the tissue by freezing. Before IHC staining slices of fixed tissue should be preincubated with specific serum to exclude unspecific antigen binding and to block the endogenous immunoglobulins.

In IHC staining the most typical marker of fibrous cap is smooth muscle cell α-actin. The antigen group specific for macrophages and/or macrophage remnants includes surface antigens such as CD68 (most common marker, observed in necrotic core of advanced plaques [76]), CD86, CD163, MARCO (scavanger receptor) and enzymes: inducible nitric oxide synthase (iNOS). Many subtypes of metalloproteinases are typical for vulnerable plaques; MMP-1, 8, 9, 13 and 14 secreted by macrophages and smooth muscle cells may be detected in necrotic core and fibrous cap area. General MMP activity could be also visualized by zymographic method, using unfixed material.

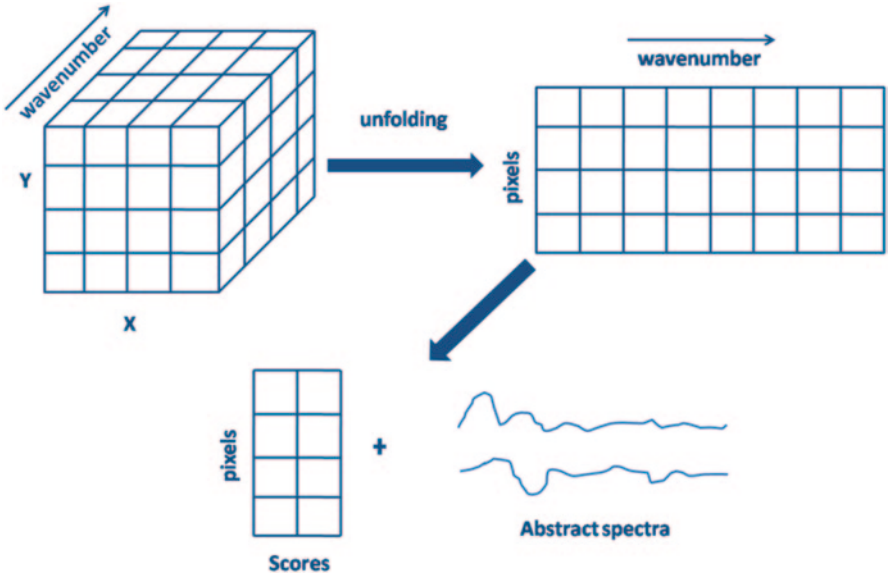


Fig. 17.4 General scheme of data analysis steps performed by different chemometric techniques. The first step is always unfolding (in case of imaging data), followed by a decomposition into abstract spectra and to scores, which correspond to "concentration" of the obtained abstract spectra in each pixel/sample

17.3 Vibrational Spectroscopy

17.3.1 Chemometrics in Spectroscopic Imaging of Atherosclerosis

Chemometrics are data analysis techniques based on mathematics, statistics and computer science. Their development and ongoing introduction into almost all branches of science is dictated by the increasing amount of information created by more modern methods [77–79]. It is this large amount of data that brings about the use of chemometrics, since in simpler cases, the scientist manages to handle the data analysis by him/herself. Therefore, with the development of high throughput imaging systems that are now available, it is a necessity to become acquainted with the capabilities of advanced data analysis techniques. Even the use of the most basic ones gives a very powerful tool to a scientist.

Before going into details of the techniques used in spectroscopy in the field of atherosclerosis, a more general description is needed. In order not to confuse the reader with different names given to different techniques, a general scheme (Fig. 17.4) is given.

The very first step in any analysis of an image is unfolding the data cube that is acquired during the measurement. In the case of measuring single spectra, this step is omitted. There exist techniques which can be performed on the full data cube

(*e.g.* PARAFAC, Tucker3), however, understanding the results can be quite complex and therefore, they have not become widely used. Therefore, it is very important to understand, that none of the methods widely applied in imaging use the spatial information and relationship between pixels—almost all of them are based only on chemical information contained in the spectra. There is currently one technique that uses that spatial relationship in infrared imaging—3D Fourier Self-Deconvolution [80] (3D-FSD). However, the principle here is to enhance the resolution of the images, rather than to gain chemical information.

At step two of the process, the data are arranged in columns by pixels or samples, with each row corresponding to a spectrum (Raman or infrared). This datablock serves as an input for further analysis.

The next step is the last step common to most of the methods, since following this step, the analysis can stop here or different parts of the results can be used for different purposes. Also, this step is performed in another way in each of the techniques and this is where differentiation starts.

17.3.1.1 Exploratory Techniques (PCA, CA)

The beginning of an analysis of an unknown dataset normally starts with a need to get a general overview of the variance present in the data. The two most important techniques are Principal Component Analysis (PCA) and Cluster Analysis (CA). Their use is also widespread across all scientific disciplines and therefore their understanding is extremely crucial. PCA is a Factor Analysis (FA) technique and sometimes the naming is interchangeable, but it needs to be pointed out that PCA is always a Factor Analysis technique, but there are other FA techniques than PCA.

PCA decomposes the data into abstract spectra (called loadings in this technique) according to two conditions—covering the maximum variance with a loading and each successive loading must be orthogonal to all of the previous. This causes that usually a few loadings are enough to cover most of the variance in a dataset. It is the very power of PCA that in a fast way it can find the main source of variation in analyzed data and to possibly give it a chemical meaning by comparing the abstract spectra. If the scores on the samples corresponding to a single loading are plotted against another loading's scores, a hidden structure of the data can sometimes be revealed. In this way a possible grouping inside the dataset can be found, which is another important advantage of PCA. Unfortunately, the grouping may not always appear, especially not in the way the user expects it to happen, *e.g.* to discriminate control samples from ones influenced by a factor. This is one of the disadvantages of PCA—it follows the biggest sources of variance, not necessarily the ones the user is interested in. Moreover, when dealing with spectra, it is very common to have big variance sources coming from measurements, which sometimes can totally dominate the analysis. Among such sources are very common scattering effects in FT-IR. In Raman spectroscopy it can be the intensity of the signal, which is dependent on the confocality of the microscope. This means that by using a confocal system the data is acquired only from a confocal surface and, in the case of biological

samples, this may cause a large variation of the Raman signal when reaching the border of the sample or a denser area.

Cluster Analysis is a group of techniques used for comparing pixels/samples based on their similarity or dissimilarity. This is achieved by calculating a so called distance (which is a measure of similarity) between two pixels/samples. The distance can be calculated currently in around 70 ways, but the ones most widely used in spectroscopy are: Euclidean, Manhattan, Mahalonobis and based on Pearson's correlation coefficient (e.g. D-Values). The algorithm calculates a distance matrix, which contains relationships between all pixels/samples present in the dataset, by comparing the spectra using a distance measure. Then, based on the relationships from distance matrix, a linkage is performed, which joins together pixels/samples of closest values of distance measure forming clusters. By defining the number of clusters the user determines the number of classes to which the dataset is split. In the end, each pixel/sample is attributed a class membership value and these values can be then used to create an image, in which each class is assigned a different color. The safest version of CA is Hierarchical Cluster Analysis (HCA), which uses all of the pixel/sample relationships to create a full hierarchy of the samples with regard to their similarity. However, since it uses all of that information it is very time demanding to calculate. Therefore, simplified versions have been invented, out of which the K-Means Clustering (KMC) is most widely spread. In this algorithm, a defined K-number of clusters are assumed in the analysis—in HCA the analysis is done without any assumption and only after the user chooses the number of classes. During the analysis only the distances of the samples to cluster means are calculated and based on that assigned to the closest cluster. The process is iterative and in the first step it gives random positions to the K-clusters. After each iteration the position of the cluster mean is recalculated and the sample distances are corrected. The algorithm repeats each step 'moving' the K-clusters means until in successive steps the positions remain sufficiently the same. Therefore, the analysis may not be (and often is not) reproducible, since it may fall into local minima—this is the biggest disadvantage of the K-Means Clustering (KMC). However, due to strongly enhanced computational efficiency it is a very popular method.

17.3.1.2 Classification (LDA, M-LDA, PCA)

The predominant aim of any classification technique is to find a mathematical method of distinguishing samples of different classes in an unknown dataset. The assignment is based on previous training of the model using a Training dataset. Usually, the training classes are obtained from exploratory analysis but may be as well imposed by the user. The second situation is more common in biomedical research, since it is often required to assign samples based on their spectra to a predefined set of classes, e.g. different stages of atherosclerosis progression.

In terms of mathematical treatment this idea is most commonly performed on the scores associated with abstract spectra (Fig. 17.5). In a simple space with two abstract spectra the idea is to find a line, to which new unknown sample scores can

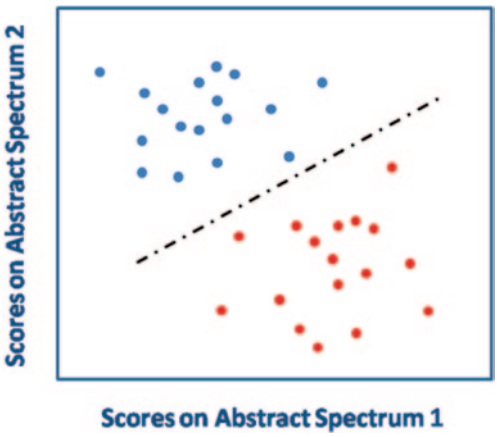


Fig. 17.5 General scheme of a classification problem

be compared and then a decision is made if the sample belongs to red or blue class. This is not only the basis of Linear Discriminant Analysis (LDA), but of most of classification problems. LDA can work on any kind of scores provided from the dataset (e.g. Mahalonobis distance) and in a special case on Principal Components obtained in PCA, giving the final name of LDA-PCA. In more advanced modeling schemes the function is not necessary linear, but these approaches require high level of careful data treatment, however offer an increase in accuracy.

17.3.1.3 Regression and Unmixing (VCA, Linear Regression with Spectra, PLS, QITFA)

Another big question in biospectroscopy is the desire of scientists to be able to obtain not abstract spectra of components and their scores, but true concentration profiles of chemical compounds and corresponding spectra. This is in theory impossible with the large number of chemicals, each with a unique spectrum, present in a biological system. However, some main chemical groups can be defined and variation in their concentration may be of particular interest. In order to get information about concentration profiles from a mixture of spectra which form a single spectrum, one needs to 'unmix' the spectra of original constituents. This can be achieved by using spectra of reference materials known to be present in the sample and covering most of its total composition (linear regression with reference spectra) or inputting only the spectrum of a single component of biggest interest and optimize the residuals around this single spectrum (Quantitative Iterative Target Factor Analysis) [81]. When a spectrum of a compound is not present it is possible to use regression techniques such as Partial Least Squares Regression (PLSR), which need the concentration of the compound obtained from another technique to be known in several samples and then the algorithm can find an abstract spectrum corresponding to a component, whose concentration behaves in similar manner. There are techniques which are able to obtain the concentration profiles even without any

reference spectra: Independent Component Analysis (ICE) or Vertex Component Analysis (VCA). Both of which use statistical concepts to derive abstract spectra with abstract concentration profiles. This is just a short description of the very idea of the several approaches available to a spectroscopist trying to analyze his data. Interested readers are suggested to search for more specific literature.

17.3.2 Spectral Features of Atherosclerosis

17.3.2.1 Raman Spectroscopy

Classification of Types of Atherosclerotic Plaque

Raman microspectroscopy studies on atherosclerosis can be divided into separate topics depending on the type of plaque. The main goal of the majority of previous research was the identification of the atherosclerotic plaque and analysis of its biochemical composition [82]. Some of this focused on the analysis of atheroma core with the cholesterol and cholesteryl esters, as well as the analysis of the calcification at the outer base of advanced lesions. There are number of components of atherosclerosis that are easily detected by Raman spectroscopy and could be used in high-sensitivity atherosclerosis diagnosis with the use of this method. These include: collagen, cholesterol crystals, adipocytes and calcium crystals [2].

In a large number of Raman studies PCA was used to classify coronary artery tissue as either non-atherosclerotic, atherosclerotic, calcified or non-calcified plaque. This method, however, can deliver just the basic information about the plaque and does not provide the information about chemical composition of the lesion [82–84]. In most plaque samples, it was possible to correctly adjust them to the aforementioned types of plaque, when the model of Logistic regression was applied to the first and third principle components [84]. The discriminant analysis based on Mahalanobis distance, which uses distance to associate a characteristic from a sample to a specific class, is suggested to be used instead of Euclidean distance [2, 82]. The main advantage of Mahalanobis distance is that it takes into account not only the relative distance from the sample to the mean of the group, but also the covariance of the data. This results in the statistical measurement of how well the spectral parameters of a sample match the original one [2]. Independent component analysis (ICA) with the use of Mahalanobis distance gave similar results to PCA and also classifies the tissue to the general plaque groups [85].

To the group of more advanced diagnostic algorithms, which are utilized to classify plaques in coronary arteries according to their detailed, chemical composition, we can include the second linear least squares-minimalization (LSM) model [83]. It is interesting that in this method, the biochemical analysis is performed by modeling the Raman tissue spectra using a simple linear combination of the basis Raman spectra of the major components presented in the tissue. Such components include collagen type III, elastin, actin, myosin, tropomyosin, cholesterol monohydrate, cholesteryl linoleate, phosphatidyl choline, triolin, β-carotene, calcium carbonate

and calcium hydroxyapatite [86–88]. This method allows the expression of the tissue composition in relative weight percentages/relative weight fractions and provides quantitative information of such things as internal elastic lamina, collagen fibers, smooth muscle cells, adventitial fat, foam cells/necrotic core, cholesterol and carotene crystals, as well as calcification [83, 86, 87, 89, 90]. Recently, quantitative information about the amount of several components in atherosclerotic plaque was also studied with the use of the unsupervised spectral unmixing algorithm, vertex component analysis (VCA) [69].

Characteristic bands for the different morphological and biochemical features present in atherosclerotic plaque with the corresponding compounds assignment and their modes are presented in Table 17.1. Figure 17.6 presents an example of a chemical image of the cross-sections of a brachiocephalic artery with plaque obtained through the use of the Raman spectroscopy. The results of K-means Clustering (KMC, see Chap. 3.1.1.) represent the morphology of the sample and without staining, allowed us to recognize such features as remodeled media, heme, internal elastic lamina/fibrous cap, lipids, calcification and calcium nodule.

The Raman maps were recorded using a WITec confocal CRM alpha 300 Raman microscope equipped with an air cooled solid state laser operating at 532 nm and a CCD detector which was cooled to −82ºC. The laser was coupled to the microscope via a single mode optical fibre with a diameter of 50 μm. A dry Olympus MPLAN (100/0.90NA) objective was used. Spatial resolution was equal to 1 μm.

The presence of the lipid core is evidenced by bands connected with the cholesterol fatty acid esters and cholesterol. These bands are observed mainly at around 1440, 1670 and 1740 cm^{-1} (yellow class—lipids). The bands at 960 cm^{-1} and 1072 cm^{-1} are characteristic for calcified plaque and are assigned to the symmetric stretching vibration of phosphate groups and the combination of the antisymmetric stretching vibration of phosphate groups with the symmetric stretching vibration of carbonate groups, respectively (calcium nodule class). The gray class represents the lipid core with calcification features, because the average spectrum of this class corresponds to hydroxyapatite and lipids. The internal and external elastic laminas contain both elastin and collagen, but with a higher concentration of elastin. It was shown from Raman studies that elastic lamina is very consistent within and between coronary artery samples [87]. The Raman spectra of the collagen fiber and fibrous cap differ from elastic lamina mainly with the absence/decrease of the bands connected with desmosine and isodesmosine, which are typical for elastin. Moreover we observe the appearance/increase of the band connected with hydroxyproline, which is typical for collagen. In Fig. 17.6, elastic lamina and fibrous cap are illustrated as a green class. The Raman spectra of tunica media, which is built mainly from muscle cells, originates mainly from the protein bands feature. The amide I at 1664 cm^{-1}, amide III near 1257 cm^{-1} and phenylalanine at 1004 cm^{-1} are the main bands of this class. In Fig. 17.6, this class is called remodeled media, as it corresponds to the pathologically altered tissue. The average spectrum of the heme class has the strongest marker bands at 745, 1130 and 1580 cm^{-1}, which are connected with the vibrations of the porphyrin ring when the excitation laser wavelength is equal to 532 nm. This class is also interesting to image, as heme features

Table 17.1 The most prominent bands present in atherosclerotic tissue Raman spectra with their modes, corresponding compound assignment, the aorta/plaque part with the highest concentration of the compound and references

Wavenumber [cm^{-1}]	Mode assignment	Compound assignment	Aorta/plaque part	Ref.
430, 544–559	Steroid rings stretching	Crystalline cholesterol—Unesterfied (free) cholesterol	Necrotic core/Steroid nucleus	[92–94]
702	Steroid rings stretching	Crystalline cholesterol—Unesterfied (free) cholesterol and Cholesteryl esters	Necrotic core/Steroid nucleus	[92–94]
855	CH_2 twisting	Proline and hydroxyproline in collagen	Collagen fiber and fibrous cap	[2, 95, 96]
878	Steroid rings stretching	Crystalline cholesterol—Unesterfied (free) cholesterol	Steroid nucleus/ foam cells/necrotic core	[92]
923	Steroid rings stretching	Crystalline cholesterol—Unesterfied (free) cholesterol	Steroid nucleus/ foam cells/necrotic core	[92]
933	C-Cstretch of the $C–COO^-$	Proline and hydroxyproline in collagen	Collagen fiber and fibrous cap	[2, 95]
940	C-C and C-H bending	Proteins	Smooth muscle cell/ media	[2]
956–962	PO_4^{3-} symmetric/ asymmetric stretching	Calcium hydroxyapatite	Large and punctate calcium mineralization	[82, 97]
957	Steroid rings stretching	Crystalline cholesterol—Unesterfied (free) cholesterol	Steroid nucleus foam cells/necrotic core	[92]
1004	Ring breathing	Phenyloalanine	Smooth muscle cell/ media (mainly)	[69]
1026	C-C and C-H bending	Cholesteryl esters	Foam cells/necrotic core	[2, 86, 92, 98]
1034	C-C and C-H bending	Proteins	Smooth muscle cell/ media	[2]
1069–1073	CO_3^{2-} symmetricstretching	Calcium carbonate	Large and punctate calcium mineralization	[2, 82, 88]
1085	C-C bending	Cholesteryl esters	Foam cells/necrotic core	[2, 86, 92, 98]
1104	Desmosine/ isodesmosine	Elastin	Elastic lamina	[2, 95]
1155–1160	C=C stretching	β-carotenoids	Yellow crystals in necrotic core	[2, 82, 86]
1176	C-C bending	Crystalline cholesterol—Unesterfied (free) cholesterol	Foam cells/necrotic core	[2, 86, 92, 98]

Table 17.1 (continued)

Wavenumber [cm^{-1}]	Mode assignment	Compound assignment	Aorta/plaque part	Ref.
1264–1274	CH_2, CH_3 bending	Crystalline cholesterol—Unesterfied (free) cholesterol	Foam cells/necrotic core	[2, 86, 92, 98]
1255–1268	Amide III, C-N stretch, α-helix conformation	Proteins mainly in elastin and collagen	Elastic lamina	[2, 82, 86, 92, 95, 99]
1301	CH_2/CH_3 deformation	Triglycerides	Fat cells/tunica adventitia adipocytes	[2, 88, 92, 100]
1328	CH_2, CH_3 bending	Crystalline cholesterol—Unesterfied (free) cholesterol	Foam cells/necrotic core	[2, 86, 92, 98]
1336	Desmosine/ isodesmosine	Elastin	Elastic lamina	[2, 95]
1440	CH_2/CH_3 deformation	Triglycerides	Fat cells/tunica adventitia adipocytes	[2, 92, 98, 100]
1443	CH_2 deformation	Crystalline cholesterol—Unesterfied (free) cholesterol and Cholesteryl esters	Foam cells/necrotic core	[2, 86, 92, 98]
1449–1452	β CH_2, CH_3	Proteins	Smooth muscle cell/ media	[2, 82, 86, 92, 99]
1517–1523	C-C/C-N stretching	β-carotenoids	Yellow crystals in necrotic core	[2, 82, 86]
1654	C=C stretching	Triglycerides	Fat cells/tunica adventitia adipocytes	[2, 92, 98, 100]
1660–1664	Amide I, C=O stretching	Proteins mainly in elastin and collagen	Elastic lamina	[2, 82, 86, 92, 95, 99]
1667–1670	C=C stretching of steroid ring	Free cholesterol and cholesteryl esters	Foam cells/necrotic core	[2, 82, 86, 100]
1668	C=C stretching	Crystalline cholesterol—Unesterfied (free) cholesterol	Foam cells/necrotic core	[2, 86, 92, 98]
1735	C=O stretching	Cholesterol esters	Foam cells/necrotic core	[2, 86, 92, 98]
1747	C=O stretching	esters	Fat cells/tunica adventitia adipocytes	[2]
2850	CH_2symmetric stretching	Cholesterol esters	Foam cells/necrotic core	[93]

Table 17.1 (continued)

Wavenumber [cm^{-1}]	Mode assignment	Compound assignment	Aorta/plaque part	Ref.
2870	CH_3symmetric stretching	Crystalline cholesterol—Unesterfied (free) cholesterol	Foam cells/necrotic core	[93, 101]
2885	Fermi resonance CH_2stretching	Cholesterol esters	Foam cells/necrotic core	[93]
2933–2945	symmetric CH_3/assymmetricCH_2/out-of-plane chain	Crystalline cholesterol—Unesterfied (free) cholesterol	Foam cells/necrotic core	[93, 101]

like intraplaque hemorrhage and thrombotic plaques (plaques exhibiting a luminal thrombus) are considered as having a huge impact on plaque vulnerability (see Chap. 3.2.1.2.) [47, 86].

In Search of vulnerable plaque spectroscopic features

One of the most important goals of Raman spectroscopy is the identification and detection of unstable, vulnerable plaque. In order to do this, the morphological markers of vulnerable atherosclerotic plaque should be described [102]. Both IR [73, 103] and Raman [47, 86, 104, 105] spectroscopies were developed for studying the composition of plaque and detecting vulnerable plaque. As shown previously, the chemical composition along with the morphology can provide more information about plaque stability than the degree of stenosis [87, 106]. Therefore, it seems that the combination of vibrational spectroscopy with staining should be the perfect solution to this problem.

The most basic statement claims that unstable plaque is defined as a thin fibrous cap overlying a large pool of necrotic lipid-rich material with high risk of rupture [2, 73, 82, 87, 105, 107, 108] (see Chap. 4.2.). It has been shown that such features as intraplaque hemorrhages and thrombotic plaques (plaques exhibiting a luminal thrombus) are also considered to have a considerable impact on plaque vulnerability [73, 109]. Other possible observations are the activated macrophages near the fibrous cap, inflammation, reduced collagen synthesis, local overexpression of collagenase and smooth muscle cell apoptosis [105].

Such a description clears up why the evaluation of vulnerable plaque was not encountered during *in vitro* studies. This is explained by the fact that some species, like intraplaque hemorrhagic and hemoglobin concentration in plaque, have a huge role in defining vulnerable plaque and can be correctly observed only *in vivo* [47].

Raman spectroscopy allows for the detection not only of the cholesterol and calcium mineralization, but [86, 110] also such morphologic features as fibrous cap [111], necrotic core/foam cells [73], ceroid presence and its formation [112, 113], hemoglobin (Hb), β-carotene (not present in all plaques [114]) and thrombus on the surface of plaque [73, 105]. The ability to detect this wide array of features makes

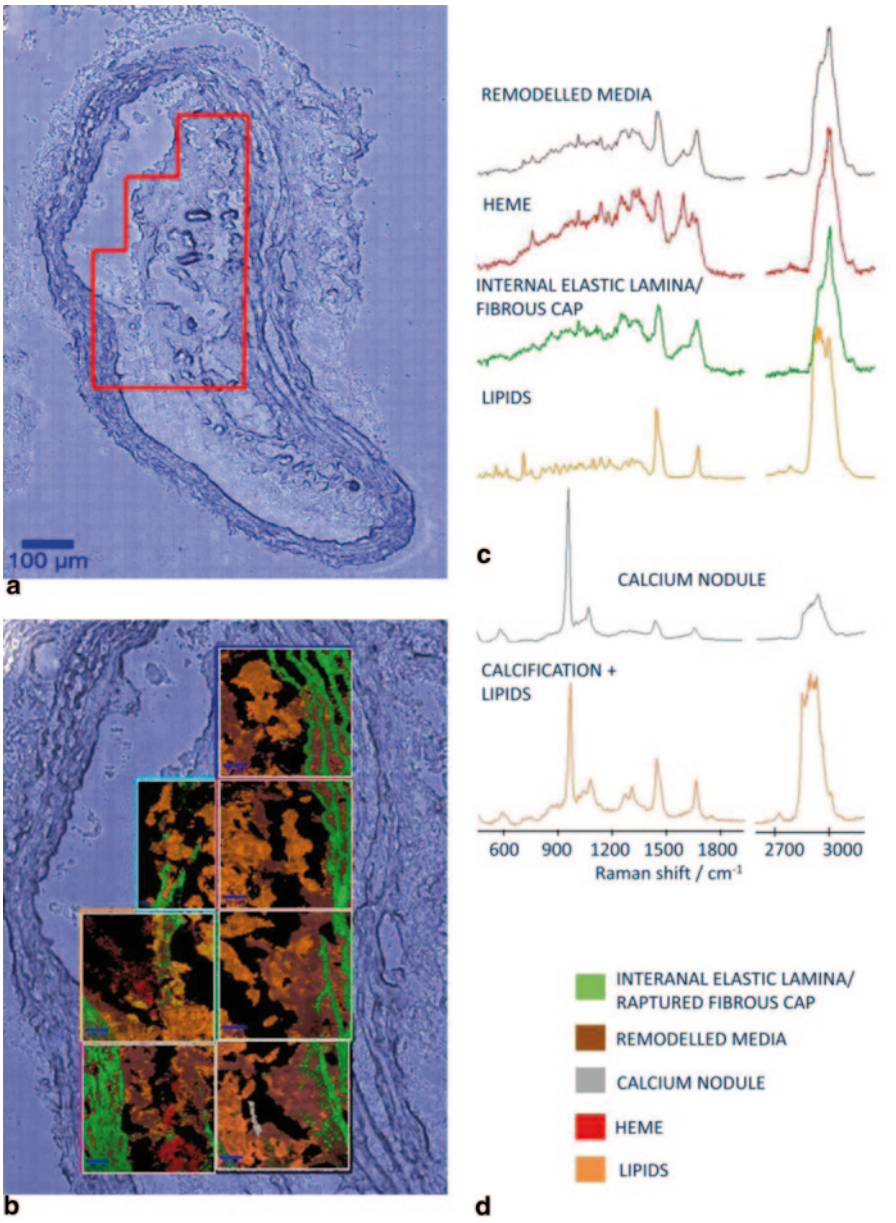


Fig. 17.6 **a** A visual image focused on the cross-sections of a brachiocephalic artery (BCA) coming from a 6-month old ApoE/LDLR−/− mouse fed LCHP diet [91]. The measured part of BCA tissue is presented in the red frame. **b** The K-means Clustering (KMC) results with the five main classes which were chosen artificially to represent the morphology of the sample. Classes include remodeled media, heme, internal elastic lamina/fibrous cap, lipids and calcium nodule. **c** The average spectra of the respective 5 classes. The spectrum which represents the pure calcification nodule (*gray spectrum*) was distinguished from the class which originates from calcification and lipids (*bottom, yellow spectrum*). **d** Code color for presented in the figure classes

this imaging technique a very powerful method for the detection of unstable plaque. Previous Raman studies also revealed the possible presence and pathophysiological importance of retinal palmitate, ascorbic acid in ceroid deposits and α-tocopherol [87]. It is important to mention that foam cells and necrotic cores are difficult to recognize with the use of Raman spectroscopy [2, 114].

The combination of Raman techniques with virtual histology intravascular ultrasounds, or morphology-based diagnosis, allows for not only the definition of the structures of plaque, but also relative quantification of some components—mainly lipids and calcium salts, but also elastin, collagen, actin, triolein, cholesterol linoleate, free cholesterol and β-carotene [86–88, 92, 95, 105, 110].

One of the most important parts of a lesion, which is able to provide the information about the plaque vulnerability, is the lipid fraction. Such fractions can be divided into two groups. The first, cholesteryl esters, exists mainly in the liquid state, while the second is the group of free cholesterol, which exists in the crystalline form [100, 101, 115]. It has been found that the presence of the cholesteryl esters may soften the plaque, while the presence of cholesterol (crystalline form) makes it more stable [87, 100, 101, 116]. On the other hand, there were studies, which show that the presence of cholesterol crystals induces inflammation and is connected with plaque rupture [117, 118], presence of macrophages [119], as well as acute coronary deaths [118, 120]. It was also found that the calcified deposits are located within the area of cholesterol accumulation [88].

Figure 17.7 shows an example of cholesterol crystals observed inside plaque with the Raman integration map and the spectrum of the cholesterol crystal and cholesteryl esters obtained from the atherosclerotic plaque. Two main bands at 1667–1674 cm^{-1} (C=C stretching) and 1440 cm^{-1} (C-H bending) are shared for both free cholesterol and cholesteryl esters, while the band at 1655 cm^{-1} is connected with C=C stretching in unsaturated fatty acids [100]. In the Raman spectrum of cholesteryl esters, bands observed at lower frequencies due to cholesterol moiety, (connected mainly with the vibration modes of steroid rings) are not as distinct as in the case of pure cholesterol. The spectra of cholesteryl esters are mainly characterized by bands originating from a hydrocarbon chain, the C=O stretching at around 1740 cm^{-1} and the presence of the strongest band in the C-H stretching region at 2885 cm^{-1}. Spectra characteristic for crystalline forms of cholesterol (anhydrous and monohydrate cholesterol) originate mainly form cholesterol moiety (steroid rings) and are observed at around 550, 605, 700, 742 and 850 cm^{-1}. Raman spectroscopy allows for the not only the recognition of cholesteryl esters from cholesterol, but also the detection of the specific type of crystalline cholesterol (anhydrous, monohydrate, hydroxycholesterol) [94, 100]. Near-infrared Raman spectroscopy has also been used in quantitative analysis of cholesterol and cholesteryl esters in human atherosclerotic plaques [121]. It was found that for quantitative analysis of cholesterol and cholesteryl esters in human atherosclerosis plaques, the most powerful method was partial least squares regression (PLS) based on factorial analysis [100, 122].

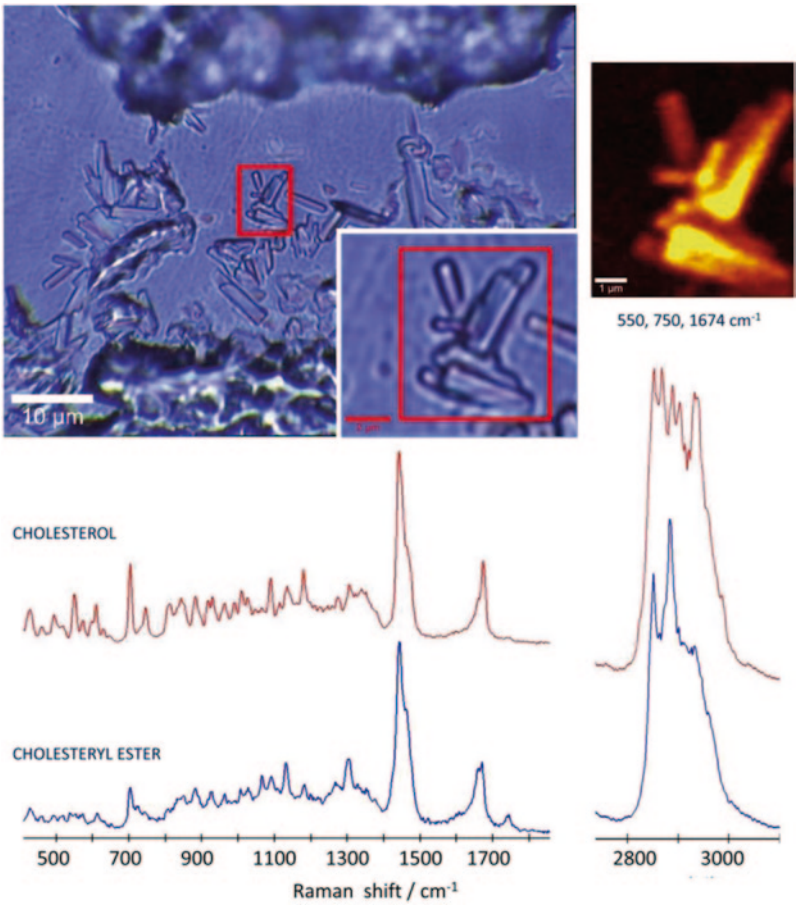


Fig. 17.7 A microphotograph of the lipid core of a brachiocephalic artery (BCA) taken from a 6-month old ApoE/LDLR−/− mouse fed LCHP diet [91]. The measured part of lipid core focused on the crystals of cholesterol and is presented in the red rectangle. The same integration maps were obtained for characteristic bands of cholesterol crystal (550, 750 and 1674 cm^{-1}). The yellow color represents the highest relative concentration of the cholesterol. The spectra of cholesterol (cholesterol crystal) and cholesteryl esters (different than the crystal part of the lipid core) were recorded using a WITec confocal CRM alpha 300 Raman microscope equipped with an air cooled solid state laser operating at 532 nm and a CCD detector which was cooled to −82 °C. The laser was coupled to the microscope via a single mode optical fibre with a diameter of 50 µm. A dry Olympus MPLAN (100/0.90NA) objective was used. 120 scans were collected and averaged with the integration time equal to 0.3 s

In Vivo Studies of Atherosclerotic Plaques

All of the examples mentioned above, together with the possibility of the application of optical fiber technology in clinical Raman instruments, make Raman spectroscopy a suitable method for not only *in vitro,* but also *in vivo* analysis of athero-

sclerotic plaque [23, 47, 123–125]. The first real-time *in vivo* studies with the use of Raman spectroscopy of human atherosclerosis were carried out in 2006 [47]. In that work, 74 sites during caroid endarterectomy and femoral artery bypass surgeries were examined with the sensitivity and specificity of 79 and 85 %, respectively. However, it is important to mention that such preliminary studies require long collection time to obtain acceptable signal-to-noise ratios [47]. In most *in vivo* experiments on sheep and humans, aortas were saline-flushed while spectra were acquired [110, 123]. Recently, C. Matthäus et al., employed Raman-probe spectroscopy to characterize *in vivo* the plaque compositions in arterial walls of a rabbit model and were able to analyze the biochemical composition of atherosclerotic plaques without markers, in a non-destructive way [126].

With the use of quantitative *in vivo* Raman spectroscopy, it was possible to create the scoring criteria of vulnerable plaque in human atherosclerosis (with a high probability of rupture) [47]. To these criteria, we can include mainly thrombosis, ulceration, fibrous cap, hemoglobin, calcification, collagen and the sum of the contribution of cholesterol and foam cells/necrotic core. Multimodal spectroscopy (where the combination of DRS, IFS and Raman was applied) is also used in *in vivo* studies of the human vulnerable atherosclerotic plaque during femoral bypass and carotid endarterectomy surgeries [73]. This method allows for the ability to detect vulnerable plaque with a sensitivity of 96 % and specificity of 72 % and the possibility to detect the thrombus on the surface of plaque [73]. Non-linear techniques, such as coherent anti-Stokes Raman scattering (CARS), have also been suggested to be effective in the study of atherosclerotic plaque [127].

17.3.2.2 IR Spectroscopy

IR Spectral Features of Atherosclerotic Tissue

FT-IR analysis of the atherosclerotic tissue is mostly concentrated on lipids and proteins, since those chemical classes are easily distinguishable in infrared spectroscopy. Proteins have characteristic amide bands (I, II A and B—1655, 1550, 3300 and 3100 cm^{-1}, respectively) [53, 54], which are quite specific and usually do not overlap with other tissue constituents. Unique profiles of amide I and II bands allow to estimate the content of secondary structures of proteins and this allowed to distinguish between collagen of type I, III and elastin [60].

Lipids on the other hand have a specific carbonyl band at 1740 cm^{-1}, which is usually assigned to cholesteryl esters but also to triglycerides. However, a more general analysis of lipid standards [67] revealed that it is not possible to distinguish between triglycerides and cholesteryl esters based only on the 1740 cm^{-1} band shift—cholesteryl esters tend to have slightly lower wavenumbers around 1735 cm^{-1}, while triglycerides in the 1740–1750 cm^{-1} range. This is due to the fact that the C=O band position in triglycerides depends strongly on the conformation of the three chains and therefore also on the exact state in which they are present. Different polymorphs show large shifts of this band.

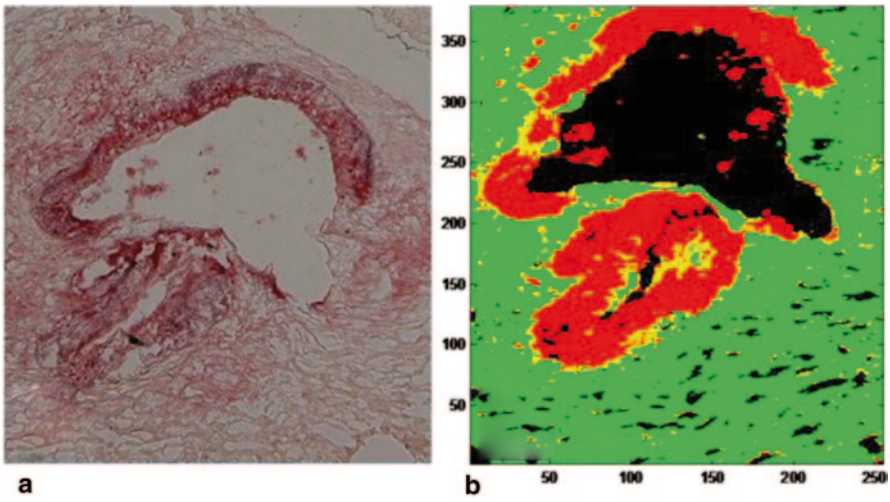


Fig. 17.8 (a) Microphotograph of Oil Red O stained cross-section of an aorta and a (b) HCA result of a FT-IR image of adjacent cross-section, with atherosclerotic plaque class colored red. Both tissues were measured on normal glass and HCA was performed with D-Values in the 2800–3025 cm^{-1} range

The CH_2/CH_3 vibration band at 2800–3025 cm^{-1} is present for both proteins and lipids and therefore was not used for any discrimination, however, the shape of the band is quite different for lipidic long chains than for average protein present in the tissue. Moreover, in this region the fatty chain unsaturation level can be determined [64] by a feature at 3008 cm^{-1}. It was suggested [67] that for cholesteryl esters with an increase in the number of double bonds in the fatty acid moiety the band shifts from 3003 to 3009 cm^{-1} for cholesteryl oleate and cholesteryl linoleate, respectively.

Pure cholesterol was identified as having a quite specific (with respect to other lipids in the tissue) band at 1055 cm^{-1}. Unfortunately, this band is present in a region rich in vibrations (1000–1100 cm^{-1}) of many constituents, including sugar moieties and PO^{2-} containing compounds. However, inside the atherosclerotic plaque cholesterol predominates in concentration and this band can be successfully used to indicate cholesterol presence.

Calcification is more pronounced in Raman spectra but still in FT-IR it is possible to find corresponding, yet quite broad bands [57, 60], i.e. ~3400, ~1000–1080 and ~600 cm^{-1}. These spectral regions are common to many compounds that are present in the tissue (as mentioned earlier) and therefore FT-IR analysis of calcification is of limited applicability and interest. However, if FT-IR is used not as a standalone technique but is correlated with, e.g. nanoindentation, it can give interesting results [66]. Using band's integrations of 3 main components (lipids, protein matrix and mineral content) vs. modulus from ATR measurements, Ebenstein et al. [66] showed that Mineral:Matrix ratio correlates well with increasing modulus, which is not surprising, however they've also found an indication that Lipid:Matrix ratio may also correlate.

Table 17.2 The most prominent bands present in atherosclerotic plaque of infrared spectra with their mode and corresponding compound assignment [67, 127].

Wavenumber [cm^{-1}]	Mode assignment	Compound assignment
1041	C-O stretching	Oligosaccharides and polysaccharides
1055		Cholesterol
1080–1095	PO_2^- symmetric stretching, C-O stretching	Phospholipids and nucleic acids glycogen, oligosaccharides and glycolipids, calcification
1151	C-O stretching	Glycogen and nucleic acids
1170	CO-O-C asymmetric stretching of ester group	Cholesteryl esters
1203		Collagen
1239	PO_2^- asymmetric stretching	Phospholipids and nucleic acids
1395–1377	CH_2 bending	Lipids and proteins (amide III)
1460	CH_2 scissoring	Lipids
1552	C=O, NH,	Proteins (amide II)
1666	C=O, NH,	Proteins (amide I)
1733–1750	C=O stretching of ester group	Lipids and triglycerides (C=O stretching of ester group)
2850	CH_2 symmetric stretching	Lipids
2873	CH_3 symmetric stretching	Proteins mainly
2925	CH_2 asymmetric stretching	Lipids
2954	CH_3 asymmetric stretching	Lipids
3008	=CH stretching	Oleic and linoleic acid moiety
3035	olefinic ring breathing	Cholesterol
3080	OH, NH	Proteins (amide B)
3500–3200	OH, NH	Proteins (amide A), sugars, calcification

The most important spectral features of atherosclerotic plaque are given in Table 17.2.

Functional Investigation and Classification of Atherosclerotic Plaque

The very first mapping experiment concerning atherosclerotic tissue was done as early as in 1991 by Kodali et al. [54]. A point microscope with small throughput was used, therefore the results are limited to fragments of the aorta. An image of the whole cross-section was also done, but with a very large step giving a very general overview of the tissue, which yet is a nice achievement with not imaging-dedicated equipment. That work laid foundations for later research, identifying major tissue components and corresponding bands as discussed in 5.2.2.1.

The next imaging work done on a synchrotron [54] was focused on differentiation between collagen types (I and III) and elastin, which are the main components of aortic wall and fibrous cap covering the lesion. The authors managed to success-

fully find relative composition of all three proteins in each of the image pixels based on spectra of lyophilized standards. The work was a proof-of-concept for further development of FT-IR for diagnostic purposes and was done on section coming from only one patient.

In 2004 imaging equipment was sufficiently developed that, without using a synchrotron, a group led by prof. Kazarian studied fragments of aorta using transmission, macro-ATR and micro-ATR imaging [61]. All of these modes of measurements have advantages and limitations. Transmission and macro-ATR modes give a comparable size of the images (with spatial resolution in the 5–20 μm scale), while micro-ATR gives enhanced resolution of around 2–3 μm. ATR measurements require much less preparation, since penetration depth of the evanescent wave is in the 0.2–2 μm range and therefore tissue of any size can be placed and measured, while for transmission mode a cross-section of approximately 10 μm is required. The only drawback of ATR measurements is the destructivity of soft tissues, which is often omitted in discussions, but is still present [1].

The large scale imaging allowed to image the *tunica media* and cholesteryl esters, claiming a quite uniform distribution of the esters. However, when a micro-ATR mode was used it became evident that cholesteryl esters form clusters in a surrounding matrix. Generally, a high level of heterogeneity was reported.

The heterogeneity was furthermore confirmed by Wrobel et al. [67] in 2011 with the FT-IR transreflection imaging supported by Hierarchical Clustering Analysis. It was shown that by using HCA it was possible to find two atherosclerotic plaque classes, which after subtraction of their average spectra revealed differences not only in cholesteryl esters composition but also in that of pure cholesterol. This structural characterization was followed by work concentrated on protein profile in atherosclerotic tissue [68]. Again, using HCA and a distance measure based on correlation coefficient and therefore sensitive to band shape rather than intensity, it was possible to find a class localized mainly in the plaque. This analysis suggests that there are specific changes in the secondary structures of proteins inside the lesion.

Functional investigation Functional characterization can be made on many levels—one of which may be a comparison between samples of aorta with a small and a developed atherosclerotic plaque (Fig. 17.9). The classes were obtained based on HCA in 2800–3025 cm^{-1} range and the number was decided artificially to represent the morphology of the sample. In the small plaque there are only three classes, with the red one being assigned to atherosclerotic plaque. In the developed plaque however, there was a need to have 5 classes in order to fully distinguish the plaque from the vascular wall. The reason for this is that the blue class in Fig. 17.9c can be assigned as a fibrous cap, which is more similar to the wall and is not present in the small plaque. It can be seen that the amount of esters (based on the band around 1735 cm^{-1}) is comparable in atherosclerotic classes in both samples, but the difference in the cholesterol band (1055 cm^{-1}) quite big.

An interesting functional characterization of atherosclerotic plaques has been performed with a series of 3 articles by prof. Kazarian group. The first one [64]

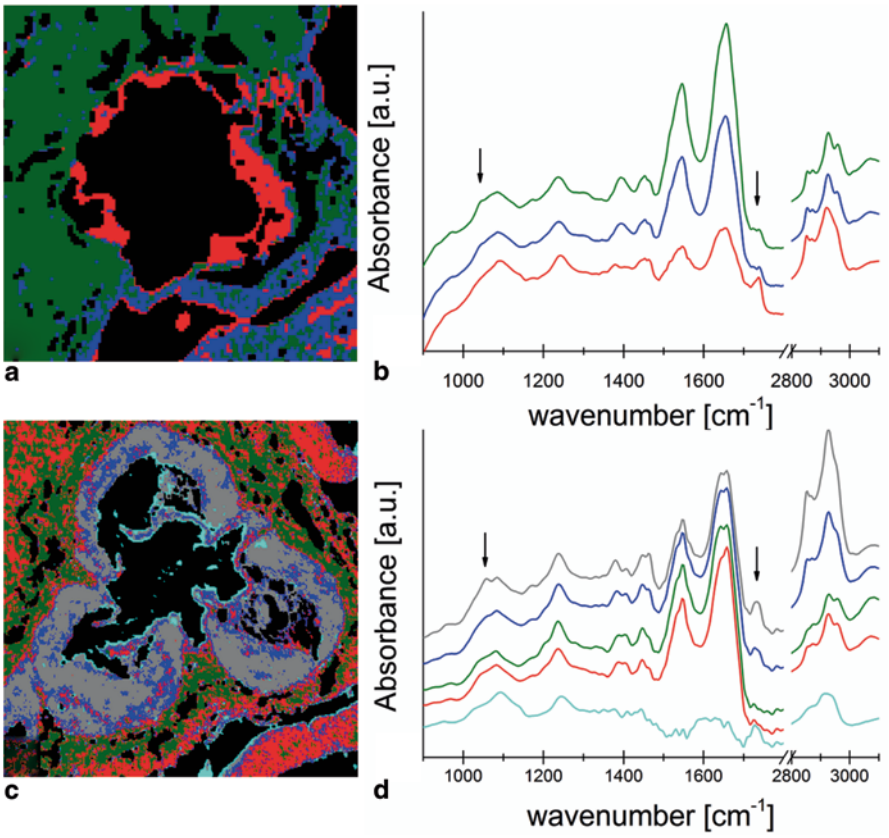


Fig. 17.9 Comparison of HCA results of two cross-sections of an aorta, taken from a 2-months old (a) and 6-months old (c) ApoE/LDLR−/− mouse. The classes were chosen artificially to represent the morphology of the sample. The average spectra of the respective classes are shown (b and d) and bands of cholesterol (1055 cm^{-1}) and cholesteryl esters (1735 cm^{-1}) are indicated by arrows (spectra were offset for clarity)

is a comparison between atherosclerotic rabbits of different age (2 and 6 months) and with or without L-arginine supplement in the diet, which was confirmed to be beneficial in disease prevention as also indicated by other techniques. The semi-quantitative results show that the lipid deposits increase with age but decrease with L-arginine supplementation. Additionally, by using Principal Component Analysis (PCA) spectral profiles assigned to elastin, lipid deposits and a cryoprotectant were reported. Moreover, two different protein profiles were found in the tissues as given by PCA.

The second comparison [63] was done on transgenic mouse models: single-knockout $ApoE^{-/-}$ and double-knockout $ApoE^{-/-}iNOS^{-/-}$. The aim was to determine the influence of inducible Nitric Oxide Synthase (iNOS) on the atherosclerotic plaque composition and the structure of the wall itself. *Tunica media* was found not

to be influenced by iNOS, however, a protein nitration possibility was indicated in the plaque region. The total amount of lipids in the ApoE$^{-/-}$ strain was found to be higher, therefore showing an influence of iNOS to lesion formation. Moreover, the lipid composition balance was shifted to higher free cholesterol content in iNOS gene deficient mice. Last article describes an ingenious experiment in which diffusion of drugs into the aorta was studied [65]. A section of tissue was placed on top of macro-ATR crystal and time-resolved measurements were performed after adding a solution of benzyl nicotinate or ibuprofen. It was found that benzyl nicotinate has a higher affinity toward atherosclerotic plaque than to healthy tissue. Ibuprofen was introduced within a solution of albumin, which acted as a drug carrier, which otherwise caused ibuprofen not to enter the tissue.

Classification Raman capabilities for *in situ* measurements of the aorta endoscopically inspired a similar experiments to be performed with an ATR catheter [4, 62]. The idea is to correctly classify the whole aorta to a certain grade of atherosclerosis progression based solely on a single spectrum obtained in a single measurement. This purpose requires use of more advanced multivariate techniques such as Partial Least Squares—Discriminant Analysis (PLS-DA) or Mahalonobis Distance—Linear Discriminant Analysis (MD-LDA). Both of the approaches gave very promising results in the experiment performed on rabbit tissue [62] giving prediction accuracy of 100 %, however, no external validation set was used due to limited number of available samples. This renders the results not fully reliable, yet the experiment show the great potential of using ATR catheters beside Raman probes.

17.4 Conclusion

This chapter constitutes the review of the selected IR and Raman publications, which focus on atherosclerotic plaque analysis and presents couple of examples from our studies in this field. There is no doubts, that vibrational spectroscopy is able to distinguish morphological and biochemical changes which occurs in atherosclerotic plaques.

FT-IR spectroscopy provides information mainly about the concentration of lipids and concentration and structure of proteins. Unique profiles of amide I and II bands allow to estimate the content of secondary structures of proteins, as well as distinguish between collagen of type I, III and elastin. Different modes of IR measurements were successfully used in atherosclerotic tissues studies (transmission, macro-ATR and micro-ATR imaging) to determine such plaque features as protein profiles (mainly elastin) and lipid deposits.

Raman spectroscopy provides more specific identification of plaque contents. This method is able to distinguish such plaque features as: fibrous cap, internal elastic lamina, collagen fibers, smooth muscle cells, adventitial fat, necrotic core/foam cells, ceroid, thrombus on the surface of plaque, intraplaque hemorrhagic, carotene crystals, and calcification. Such recognition is based on the chemical differences between such groups of chemicals as: collagen, elastin, cholesteryl esters, free

cholesterol crystals, hemoglobin (Hb), calcium crystals (hydroxyapatite), ascorbic acid, α-tocopherol and β-carotene. The capabilities of Raman spectroscopy allow for the classification of the state of the aorta into such categories as non-atherosclerotic, atherosclerotic, calcified, non-calcified and vulnerable.

Analysis of results of the IR and Raman images would be impossible without the use of chemometric techniques. In order to give better insight into the composition of the plaque, different exploratory techniques, as well as different models of classification and regression, can be used. When utilized along with Raman studies, the most successful chemometric methods are principal component analysis (PCA), independent component analysis (ICA), hierarchical cluster analysis (HCA), K-means clustering (KMC), the second linear least squers-minimalization (LSM) model and the unsupervised spectral unmixing algorithm vertex component analysis (VCA). For IR images, principal component analysis (PCA) and hierarchical cluster analysis (HCA) are mainly used. It is also noteworthy that the application of advanced multivariate techniques such as Partial Least Squares—Discriminant Analysis (PLS-DA) or Mahalonobis Distance—Linear Discriminant Analysis (MD-LDA) is used to analyze IR results.

Vibrational images correlate well with histopathology, which confirms the potential of these techniques for biochemical and structural discrimination of plaque components. We believe that the future developments in these label-free techniques will provide unique high-content image of morphological and most importantly biochemical composition of the atherosclerotic plaque that may have diagnostic prognostic and therapeutic significance.

Acknowledgements This study was supported by European Union from the resources of the European Regional Development Fund under the Innovative Economy Programme (grant coordinated by JCET-UJ, No POIG.01.01.02-00-069/09). The tissues for histological and vibrational microscopy examples of atherosclerotic plaque analysis presented in this chapter were taken from ApoE/LDLR$^{-/-}$ mice [128].

References

1. Wrobel TP, Marzec KM, Majzner K, Kochan K, Bartus M, Chlopicki S, Baranska M (2012) Analyst 137:4135–4139
2. Peres MB, Silveira L, Zangaro RA, Pacheco MTT, Pasqualucci CA (2011) Lasers Med Sci 26:645–655
3. Krafft C, Sergo V (2006) Spectroscopy 20:195–218
4. Wang L, Mizaikoff B (2008) Anal Bioanal Chem 391:1641–1654
5. Kazarian SG, Chan KL (2006) BBA 1758:858–867
6. Bhargava R, Fernandez DC, Hewitt SM, Levin I (2006) BBA 1758:830–845
7. Walsh, M. J, Kajdacsy-Balla, A, Holton, S. E, Bhargava, R. *Vibrational Spectroscopy*2012, 60, 23–28
8. Reddy R, Davis B, Carney PS, Bhargava R (2011) IEEE Proceedings—International Symposium on Biomedical Imaging, 2011. doi:10.1109/ISBI.2011.5872511
9. Chew SF, Wood BR, Kanaan C, Browning J, MacGregor D, Davis ID, Cebon J, Tait BD, McNaughton D (2007) Tiss Ant 69(1), 252–258

10. Bassan P, Lee J, Sachdeva A, Pissardini J, Dorling KM, Fletcher JS, Henderson A, Gardner P (2013) Analyst 138(1):144–157
11. Diem M, Romeo M, Matthäus C, Miljkovic M, Miller M, Lasch P (2004) Infrared Phys Tech 45:331–338
12. Hughes C, Brown MD, Clarke NW, Flower KR, Gardner P (2012) Analyst 137:4720–4726
13. Heraud P, Ng ES, Caine S, Yu QC, Hirst C, Mayberry R, Bruce A, Wood BR, McNaughton D, Stanley EG, Elefanty AG (2010) Stem cell Res 4:140–147
14. Amrania H, Amrania H, McCrow AP, Matthews MR, Kazarian SG, Kuimova MK, Phillips CC (2011) Chem Sci 2:107–111
15. Stone N, Priet MC, Crow P, Uff J, Ritchie AW (2007) Anal Boanal Chem 387:1657–1668
16. Kawabata T, Mizuno T, Okazaki S, Hiramatsu M, Setoguchi T, Kikuchi H, Yamamoto M, Hiramatsu Y, Kondo K, Baba M, Ohta M, Kamiya K, Tanaka T, Sucuki S, Konno HJ (2008) Gastroenterol 43(4):283–290
17. Huang Z, McWilliams A, Lui H, Mclean D.I, Lam S, Zeng H (2003) Int J Cancer 107:1047–1052
18. Pilotto S, Pacheco MTT, Silveira L, Villaverde AB, Zangaro RA (2001) Laser Med Sci 16(1):2–9
19. Utzinger U, Heintzelman DL, Mahadevan-Jansen A, Malpica A, Follen M, Richards-Kortum R (2001) Appl Spectrosc 55(8):955–959
20. Shima MG, Song LMWK, Marcon NE, Wilson BC (2000) Photochem Photobiol 72(1):146–150
21. Ogawa M, Harada Y, Yamaoka Y, Fujita K, Yaku H, Takamatsu T (2009) Biochem Biophys Res Commun 382(2):370–374
22. Hata TR, Scholz TA, Ermakov IV, McClane RW, Khachik F, Gellermann W, Pershing LK (2000) J Invest Dermatol 115:441–448
23. Hanlon EB, Manoharan R, Koo TW, Shafer KE, Motz JT, Fitzmaurice M, Kramer JR, Itzkan I, Dasari RR, Feld MS (2000) Phys Med Biol 4(2):R1–R59
24. Mescher AL (2009) Junqueira's basic histology, twelfth edition. McGraw Hill Professional
25. Coen M, Gabbiani G, Bochaton-Piallat ML (2011) ArteriosclerThrombVasc Biol 31:2391–2396
26. Wilson FJ, Kestenbaum MG, Gibney JA, Matta S (1997) Histology Image Review: a complete illustrated review course in basic histology. Appleton & Lange
27. Camejo G, Hurt-Camejo E, Wiklund O, Bondjers G (1998) Atherosclerosis 139:205–222
28. Khalil MF, Wagner WD, Goldberg IJ (2004) Arterioscler Thromb Vasc Biol 24:2211–2218
29. Gu L, Okada Y, Clinton SK, Gerard C, Sukhova GK, Libby P, Rollins BJ (1998) Molecular cell 2:275–281
30. Haley KJ, Lilly CM, Yang JH, Feng Y, Kennedy SP, Turi TG, Thompson JF, Sukhova GH, Libby P, Lee RT (2000) Circulation 102:2185–2189
31. Mach F, Sauty A, Iarossi AS, Sukhova GK, Neote K, Libby P, Luster AD (1999) J Clin Invest 104:1041–1050
32. Weber C, Noels H (2011) Nat Med 17:1410–1422
33. Libby P (2012) Arterioscler Thromb Vasc Biol 32:2045–2051
34. Gerszten R, Tager A (2012) N Engl J Med 366:1734–1736
35. Curtiss LK (2009) N Engl J Med 360:1144–1146
36. Jones CB, Sane DC, Herrington DM (2003) Cardiovasc Res 59:812–823
37. Lindstedt KA, Kovanen PT (2004) Cardiovasc Res 15
38. Hansson GK (2005) N Engl J Med 352:1685–1695
39. Heistad DD (2003) N Engl J Med 349:2285–2287
40. Williams H, Johnson JL, Carson KGS, Jackson CL (2002) Arterioscler Thromb Vasc Biol 22:788–792
41. Bond AR, Jackson CL (2011) J Biomed Biotechnol
42. Burke AP, Virmani R, Galis, Z, Haudenschild CC, Muller JE. *J. Am. Coll. Cardiol*. 2003, *41*, 1874–1886
43. Thorp E, Tabas I (2009) J Leukoc Biol 86:1089–1095

44. Bosshart F, Utzinger U, Hess OM, Wyser J, Mueller A, Schneider J, Niederer P, Anliker M, Krayenbuehl HP (1992) Cardiovasc Res 26:620–625
45. Angheloiu GU, Arendt JT, Muller MG, Haka AS, Georgakoudi I, Motz JT, Scepanovic OR, Kuban BD, Myles J, Miller F, Podrez EA, Fitzmaurice M, Kramer JR, Feld MS (2006) Arterioscler Thromb Vascular Biol 26(7):1594–1600
46. Morguet AJ, Gabriel RE, Buchwald AB, Werner GS, Nyga R, Kreuzer H (1997) Lasers Surg Med 20:382–393
47. Motz JT, Fitzmaurice M, Miller A, Gandhi SJ, Haka AS, Galindo LH, Dasari RR, Kramer JR, Feld MS (2006) J Biomed Opt 11(2):021003, 1–9
48. Tuczu EM, Berkalp B, DeFranco AC, Ellis SG, Goormastic M, Whitlow PL, Franco I, Raymond RE, Nissen SE (1996) J Am Coll Cardiol 27:832–838
49. Vallabhasjoula V, Fuster J (1997) Nucl Med 28:1788–1796
50. Fayad ZA, Fuster V (2001) Circ Res 89(4):305–316
51. Naghavi M, Libby P, Falk E, Casscells SW, Litovsky S, Rumberger J, Badimon JJ, Stefanadis C, Moreno P, Pasterkamp G, Fayad Z, Stone PH, Waxman S, Raggi P, Madjid M, Zarrabi A, Burke A, Yuan C, Fitzgerald PJ, Siscovick DS, de Korte CL, Aikawa M, Airaksinen KE, Assmann G, Becker CR, Chesebro JH, Farb A, Galis ZS, Jackson C, Jang IK, Koenig W, Lodder RA, March K, Demirovic J, Navab M, Priori SG, Rekhter MD, Bahr R, Grundy SM, Mehran R, Colombo A, Boerwinkle E, Ballantyne C, Insull W Jr, Schwartz RS, Vogel R, Serruys PW, Hansson GK, Faxon DP, Kaul S, Drexler H, Greenland P, Muller JE, Virmani R, Ridker PM, Zipes DP, Shah PK, Willerson JT (2003) Circulation 14, 108(15):1772–1778
52. Prokoshina NA, Khovratovich NN, Khmara NF, Khatkevich MA (1985) ZhurnalPrikladnoi-Spektr 41:785–789
53. Baraga JJ, Feld MS, Rava RP (1991) Appl Spectrosc 45:709–711
54. Kodali DR, Small DM, Powell J, Krishnan N (1991) Appl Spectrosc 45(8):1310–1317
55. Manoharan R, Baraga JJ, Rava RP, Dasari RR, Fitzmaurice M, Feld MS (1993) Atherosclerosis 103:181–193
56. Hayashi J, Saito T, Aizawa K (1997) Lasers Med Sci 21:287–293
57. Li C, Ebenstein D, Xu C, Chapman J, Saloner D, Rapp J, Pruitt L (2002) J Biomed Mater Res A 64:197–206
58. Adar F, Jelicks L, Naudin C, Rousseau D, Yeh S (2003) Proceedings of SPIE 11:3–11
59. Becker A, Epple M, Müller KM, Schmitz I (2004) J Inorg Biochem 98:2032–2038
60. Wetzel DL, Post GR, Lodder RA (2005) Vib Spectr 38:53–59
61. Colley CS, Kazarian SG, Weinberg PD, Lever MJ (2004) Biopolymers 74:328–335
62. Wang L, Chapman J, Palmer RA, Alter TM, Hooper BA, van Ramm O, Mizaikoff B (2006) Appl Spectrosc 60:1121–1126
63. Palombo F, Shen, H, Benguigui LES, Kazarian SG, Upmacis RK (2009) Analyst 134:1107–1118
64. Palombo F, Cremers SG, Weinberg PD, Kazarian SG (2009) J R Soc Interface 6:669–680
65. Palombo F, Danoux CB, Weinberg PD, Kazarian SG (2009) J biomed Opt 14:044008
66. Ebenstein DM, Coughlin D, Chapman J, Li C, Pruitt L (2009) J Biomed Mater Res A 91:1028–1037
67. Wrobel TP, Mateuszuk L, Chlopicki S, Malek K, Baranska M (2011) Analyst 136:5247–5255
68. Wrobel TP, Majzner K, Baranska M (2012) Spectrochim Acta Part A: Mol Biomol Spectrosc 96:940–994
69. Lattermann A, Matthäus C, Bergner N, Beleites C, Romeike BF, Krafft C, Brehm BR, Popp J (2012) J Biophotonics 12:1–12
70. Mostaco-Guidolin LB, Sowa MG, Ridsdale A, Pegoraro AF, Smith MSD, Hewko MD, Kohlenberg EK, Schattka B, Shiomi M, Stolow A, Ko AC-T (2010) Optic Soc Am 1:59–73
71. Ko ACT, Ridsdale A, Smith MSD, Mostaco-Guidolin LB, Hewko MD, Pegoraro AF, Kohlenberg EK, Schattka B, Shiomi M, Stolow A, Sowa MG (2010) J. Biomed Opt 15(2):020501, 1–3
72. Wang HW, Langohr IM, Sturek M, Cheng JX (2009) Arterioscler Thromb Vasc Biol 29:1342–1348

73. Šćepanović OR, Fitzmaurice M, Miller A, Kong Ch-R, Volynskaya Z, Dasari RR, Kramer JR, Feld MS (2011) J Biomed Opt 16(1):011009, 1–10
74. Bancroft JD, Gamble M (eds) (2008) Theory and practice of histological techniques, 6th edition. Churchill Livingstone: Edinburgh
75. Luna LG. (eds) Manual of histologic staining methods of the armed forces institute of pathology, 3rd edition, McGraw-Hill Book Company: New York, 1968
76. Gajda M, Jawien J, Mateuszuk L, Lis GJ, Radziszewski A, Chlopicki S, Litwin JA (2008) Folia Histochem Cytobiol 46:143–146
77. Arroio A, Lima EF, Honório KM, Silva ABF (2009) Struct Chem 20:577–585
78. de Juan A, Tauler R (2003) Analyt Chim Acta 500:195–210
79. Meade AD, Byrne HJ, Lyng FM (2010) Mutat Res 704:108–114
80. Lasch P, Naumann D (2006) BBA 1758:814–829
81. Tetteh J, Mader KT, Andanson JM, McAuley WJ, Lane ME, Hadgraft J, Kazarian SG, Mitchell JC (2009) Anal Chim Acta 642:246–256
82. Silveira L, Sathaiah S, Zangaro RA, Pacheco MTT, Chavantes MC, Pasqualucci CAG (2002) Laser Surg Med 30(4):290–297
83. Bushman JH, Motz JT, Deinum G, Romer TJ, Fitzmaurice M, Kramer JR, van der Laarse A, Bruschke AV, Feld MS (2001) Cardiovasc Pathol 10:59–68
84. Deinum G, Rodriquez D, Romer TJ, Fitmaurice M, Kramer JR, Feld MS (1999) Appl Spectrosc 53:938–942
85. Silveira L, Paula AR, Pasqualucci CA, Pacheco MTT (2008) Instrum Sci Technol 36(2):134–145
86. Brennan JF III, Romer TJ, Lees RS, Tercyak AM, Kramer JR, Feld MS (1997) Circulation 96:99–105
87. Buschman HP, Deinum G, Motz JT, Fitzmaurice M, Kramer JR, Laarse A, Bruschke AV, Feld MS (2001) Cardiovasc Pathol 10:69–82
88. Faiman R (1977) Chem Phys Lipids 18:84–104
89. Romer TJ, Brennan JF, Fitzmaurice M, Feldstain ML, Deinum G, Myles JL, Kramer JR, Lees RS, Feld MS (1998) Circulation 97:878–885
90. Brennan JFB, Wang Y, Dasari RR, Feld MS (1997) Appl Spectrosc 51:201–208
91. Kostogrys RB, Franczyk-Żarów M, Maślak E, Gajda M, Mateuszuk L, Jackson CL, Chłopicki S (2012) Atherosclerosis 223(2):327–331
92. Manoharan R, Baraga JJ, Feld MS, Rava RP (1992) J Photochem Photobiol 16:211–233
93. Krafft C, Neudert L, Simat T, Salzer R (2005) SpectrochimicaActaA 61:1529–1535
94. Hawi SR, Nithipatikom K, Wohlfeil ER, Adar F, Campbell WB (1997) J Lipd Res 38:1591–1597
95. Manoharan R, Wang Y, Feld MS (1996) Spectrochim Acta Mol Spectrosc 52:215–249
96. Cárcamo JJ, Aliaga AE, Clavijo E, Garrido C, Gómez-Jeria JS, Campos-Vallette MM (2012) J Raman Spectrosc 43:750–755
97. Nelson DGA, Williamson BE (1985) Caries Res 19:113–121
98. Baraga JJ, Feld MS, Rava RP (1992) Appl Spectrosc 46:187–190
99. Diem M (1993) Intoruction to modern vibrational spectroscopy. Wiley, New York
100. Weinmann P, Jouan M, Dao NQ, Lacroix B, Groiselle C, Bonte J-P, Luc G (1998) Atherosclerosis 140:81–88
101. Small DM (1988) Arteriosclerosis 8:103–129
102. Scepanovic O, Fitzmaurice M, Gardecki JA, Angheloiu GO, Awasthi S, Motz JY, Kramer JR, Dasari RR, Feld MS (2006) J Biomed Opt 11(2):021007
103. Moreno PR, Lodder RA, Purushothaman KR, Charash WE, O'Connor WN, Muller JE (2002) Circulation 105:923–927
104. Moreno PR, Muller JE (2003) J Interevent Cardiol 16(3):243–252
105. Suh WM, Seto AH, Margey RJP, Cruz-Gonzalez I, Jang I-K (2011) Circ Cardiovasc Imaging 4:169–178
106. Burke AP, Farb A, Malcolm GT, Liang Y-H, Smial JE, Virmani R (1999) J Am Med Assoc 281:921–926

107. Woolf N, Davies MJ (1994) Sci Am Sci Med 9:38–47
108. MacNeil BD, Lowe HC, Takano M, Fuster V, Jang IK (2003) Arterioscler Thromb Vasc Biol 23(8):1333–1342
109. Virmani R, Burke AP, Kolodgie FD, Farb A (2002) J Intervent Cardiol 15(6):439–446
110. Romer TJ, Brennan JF, Puppels GJ, Zwinderman AH, van Duinen SG, van der Laarse A, van der Steen AF, Bom NA, Bruschke AV (2000) Arterioscler Thromb Vasc Biol 20:478–483
111. Alano RR, Liu CH (1994) Method for distinguishing between calcified arteriosclerotic tissue and fibrous arteriosclerotic tissue or normal cardiovascular tissue using Raman spectroscopy. US Patent, 1994, Patent No: 5293872
112. Haka AS, Kramer JR, Dasari RR, Fitzmaurice M (2011) J Biomed Opt 16(1):011011, 1–7
113. Nogueira GV, Silveira L, Martin AA, Zangaro RA, Pacheco MTT (2005) J Biomed Opt 10(3):031117/1–031117/7
114. Buschman HP, Deinum G, Motz JT, Fitzmaurice M, Romer TJ, Kramer JR, Laarse A, Bruschke AV, Feld MS (2001) Cardiovasc Pathol 10:59–68
115. Katz SS, Small DM (1980) J Biol Chem 70:9753–9759
116. Libby P, Schoenbeck U, Mach F, Selwyn AP, Ganz P (1998) Am J Med 104:14S–18S
117. Patel R, Janoudi A, Vedre A, Aziz K, Tamhane U, Rubinstein J, Abela OG, Berger K, Abela GS (2011) Arterioscler Thromb Vasc Biol 31:2007–2014
118. Abela GS (2010) J Clin Lipidol 4:156–164
119. Rajamäki K, Lappalainen J, Oörni K, Välimäki E, Matikainen S, Kovanen PT, Eklund KK (2011) Arterioscler Thromb Vasc Biol 9:2007–2014
120. Suhalim JL, Chung C-Y, Lilledahl MB, Lim RS, Levi M, Tromberg BJ, Potma EO (2012) Biophys J 102:1988–1995
121. Le Cacheux P, Menard G, Nguyen QH, Weinmann PJM, Quy DN (1996) Appl Spectrosc 50:1253–1257
122. Holton SE, Walsh MJ, Bhargava R (2011) Analyst 136 (14):2953–2958
123. Buschman HP, Marple ET, Wach ML, Bennett, B, Bakker Schut TC, Bruining HA, Bruschke AV, van der Laarse A, Puppels G (2000) J Anal Chem 72:3771–3775
124. Motz JT, Gandhi SJ, Scepanovic OR, Haka AS, Kramer JR, Dasari RR, Feld MS (2004) J Biomed Opt 43(3):542–554
125. Motz JT, Hunter, M, Galindo LH, Gardecki JA, Kramer JR, Dasari RR, Feld MS (2004) Appl Opt 43(3):542–554
126. Matthäus C, Dochow S, Bergner G, Lattermann A, Romeike BFM, Marple ET, Krafft C, Dietzek B, Brehm BR, Popp, J (2012) Anal Chem 84:7845–7851
127. Toyran N, Lasch P, Naumann D, Turan B, Severcan F (2006) Biochem J 397:427–436
128. Ishibashi S, Herz J, Maeda N, Goldstein JL, Brown MS (1994) Proc Natl Acad Sci U S A 91:4431–4435

Index

A
Atherosclerotic plaque, 506–509, 511, 512, 518, 521, 523, 526–531
  composition of, 525
  studies on, 524
  types of, 517
ab initio, 5, 67, 68, 98, 101, 113, 162, 163, 173
Aggregation, 19, 22, 25, 34, 220, 221–223, 229, 240, 402, 408
  of metallic nanoparticles, 38, 47
  of *N,N*-dimethylglycine (DMG), 187
  proteins, 221, 453, 458
Amino acids, 84–89, 91, 92, 119, 163, 180, 184, 189, 257, 329, 332, 333, 337, 344, 348, 351, 365
  role in organisms, 85
  SERS studies of, 38, 39
  solubility, 89, 90
  studies on, 101, 105–108, 123, 128, 134, 137, 138
Anthraquinones, 259, 261, 281, 282, 301
Apparent inverse order, 240
Atherosclerosis, 9–11, 506–509, 517, 530
  spectroscopic imaging of, 513
  Raman microspectroscopy studies on, 517, 525
Attenuated total reflection (ATR), 5, 218, 285, 432–435, 485, 486, 526, 528
Autopower spectrum, 236, 237, 241

B
Band
  Amide I, 119, 192, 219, 225, 226, 230, 257, 323, 361, 364, 375, 421, 428, 432, 448, 454, 457, 489, 530
  Amide I', 225, 226
  Amide II, 120, 192, 224, 424, 428, 432, 489
  Amide II', 225
  Amide B, 224, 230
  C=O stretch, 230, 461
Bayesian classification, 497
Biological application, 5, 35, 139, 174
Biomedical application, 7, 8, 199, 357, 392, 403, 415, 475, 476, 486
Biologically Relevant Molecules, 179–201, 216
Biomarker, 9, 45, 324, 367, 371, 373, 401–404, 413, 415, 420, 421, 424, 446, 447, 450, 498
Biosensors, 44, 48
Biotin, 108, 225

C
Cancer pathology, 357, 475
Carbohydrates, 6, 8, 64, 65, 67, 258, 269, 323, 356, 382, 421, 426, 450, 462, 488
Carotenoids, 6, 8, 226, 266–271, 281, 282, 290–298, 301, 310–324
Cardiovascular disease, 300
Cellular analysis, 506
Cervical cancer, 357, 358, 362–364, 367–372, 428, 446, 450, 451
Chemometrics, 7, 9, 15–25, 75, 216, 241, 257, 259, 279–284, 286–291, 297–303, 310, 311, 314, 319, 323, 361, 441, 497, 513, 531
Chemical imaging, 479
Chemotherapeutics, 357, 373, 377–379
Chemotype classification, 7, 267
Chirality, 38, 39, 64, 84, 125, 170
Circularly polarized luminscence, 162, 173

M. Baranska (ed.), *Optical Spectroscopy and Computational Methods in Biology and Medicine,* Challenges and Advances in Computational Chemistry and Physics 14,
DOI 10.1007/978-94-007-7832-0, © Springer Science+Business Media Dordrecht 2014

Combination band, 100, 109, 112, 117, 137, 223, 230, 235, 300
Conformation
α-helix, 74, 75
PPII
β-sheet, 74, 424
β-strand, 424
β-turn, 424
Correlation, 10, 19, 74, 91, 94–96, 108, 109, 163, 213–216, 223, 235, 238, 262, 294, 331, 340, 347, 369, 420, 462
Asynchronous, 74, 109, 137, 214–216, 222–229, 241, 299
Generalized 2D, 214, 216, 238
Heterospectral, 74, 229, 233–235
Homospectral, 234
Moving window 2D, 236
Synchronous, 74, 109, 137, 214–216, 219–223, 235–238, 261, 299
bν-correlation
kν-correlation
Coumarins, 183, 204
Cross peak, 109, 235, 261

D
2D correlation spectroscopy (2D-COS), 108, 214, 216, 219, 223, 240, 289, 299
Data clustering techniques, 494
Data processing, 294, 393, 421, 465, 466
DFT, 38, 67–69, 72, 96, 106, 113, 163, 262, 267, 274, 311
Discrete frequency, 84
Disease diagnostics, 355–393

E
Early diagnosis, 10
Electronic Circular Dichroism (ECD), 4, 76, 93, 161–175, 217
Experimental, 61, 71, 74, 91, 94, 106, 174

F
Flavonoids, 258, 259, 261, 267, 281–283, 297–300
Fourier transform infrared imaging theory, 94, 195, 227–229, 298, 302, 356, 357, 363, 364, 371, 419–465, 498, 506

H
Heart disease, 507
Hydrogen bonding, 125, 126, 187, 191, 216, 224

I
Imidazoacridines, 22
Immunoassay, 45, 401–415
Indigoids, 259, 283
Induced chirality, 38, 39, 64, 84
Infrared Spectroscopy, 8–10, 183, 184, 191, 193, 214, 215, 217, 254, 256, 285, 294, 356, 357, 364
Instrumentation, 64, 182, 378, 420, 421, 430, 460, 475
Intracellular SERS, 47, 48

L
Lasers, 30, 45, 284, 291
Lipid bilayers, 22, 225, 239
Lipid
Dimyristoylphosphatidylglycerol (DMPG), 225
Liver diseases, 9, 463
Low Temperature Inert Matrices, 179–207

M
Mapping and imaging, 431, 435
Matrix isolation, 3, 4, 94, 101, 123, 181–184, 187, 188, 191–193
Metabolic fingerprinting, 253, 254
Metrics, 491–493
Microfluidics, 402, 412–415
Multivariate curve resolution (MCR), 20, 241, 316
Multivariate statistical analysis, 331

N
Nanoparticle probes, 36, 47, 50, 51,
Natural dyes, 259, 262, 279, 281–303
Neurodegenerative pathologies
Noda's rule, 215, 216, 241
Nucleic acid bases, 180, 191–193
Nucleic acid structure, 64, 170

O
Overtone, 100, 194, 223, 234, 286, 289, 300

P
Perturbation
concentration, 233
H/D exchange, 225
pH, 228, 229
pressure, 227
temperature, 219–223
Phase angle, 238, 443
Phenolic compounds, 6, 183, 199, 258, 269, 300

Poly(L-glutamic acid) (PLG), 235, 237
Poly(L-lysine) (PLL), 227, 237
Porcine plasma, 222, 229
Pretreatment, 74, 216, 217–219, 224, 279, 287
Primary and secondary metabolites, 253
Principal component analysis (PCA)
  loadings, 18–20, 23, 75, 240
  scores, 24, 226, 289
Process
  adsorption, 32, 38, 39, 40, 44, 224, 403
  aggregation, 19, 22, 24, 34, 38, 47, 187, 190, 220–223, 229, 240, 402, 408, 453, 458
  association, 90, 224
  denaturation, 75, 166, 168, 171, 220–224, 229, 232, 375
  fibrillation, 215, 221, 223, 240
  folding, 93, 161, 167, 168, 181, 215, 218, 219, 240
  oligomerization, 221
  unfolding, 75,168, 171, 219–224, 229, 231
Prostate breast, histopathology, 9, 362, 477, 496
Protein structure, 75, 84, 90, 164, 215, 220, 223, 225, 228, 233, 257, 461
Protein
  Amyloid, 166, 169, 175, 215, 221, 454–457
  Bovine pancreatic trypsin inhibitor (BPTI), 227
  Bovine serum albumin, 44, 228, 405
  Centrin, 227
  Chicken annexin A5, 220
  Cytochrome c (cyt c), 40, 221, 222
  Epidermal growth factor receptor (EGFR)
  Ferricytochrome c (fcyt c), 221
  Hemoglobin, 221, 222, 509, 521, 525, 531
  Hepatitis C virus core protein (HCVc-120), 226
  Human procarboxypeptidase A2 (ADA2h), 221
  Human serum albumin (HSA), 224, 225, 228
  Immunoglobulin (IgG), 224, 226, 402, 512
  α-lactalbumin (α-la), 168
  β-lactoglobulin (β-lg), 231, 233
  Lipoxygenase (LOX), 227
  Melittin, 227
  Methionine adenosyltransferase (MAT), 220
  Myoglobin (Mb), 221, 222, 225
  Ovalbumin, 223, 224
β-purothionin

R

RibonucleaseA(RNase A), 220–222, 463
  Sarcoplasmic reticulum Ca2+-ATPase, 220
  Streptavidin, 225
  Quality, 16, 67, 71, 94, 99, 137, 163, 218, 264, 268, 286, 350
  Quantum-chemical calculations, 93, 311–314
Radiation damage, 4, 5, 35, 192

S

Secondary structure, 75, 92, 137, 164–170, 215, 219–225, 228, 229, 240, 424, 453, 458, 525, 528, 530
Sequence of intensity changes, 238
SERS of biological molecules, 47
Silk fiber, 229
Similarity Analysis, 330, 337, 350–352
Spectroscopy
  Fluorescence, 138, 169, 173, 220, 228, 264, 284, 312, 316, 357, 388, 391
  Infrared, 3, 7–10, 183, 184, 214, 215, 355, 364, 369, 379, 381, 420
  Mid-IR, 106, 109, 194, 195, 217, 223, 224, 228, 229, 286, 299, 451, 482, 484
  Near-IR (NIR), 9, 109, 194–196, 217, 223, 235, 241, 441
  Raman imaging, 4, 6–10, 30, 47, 63–76, 97–99, 161, 229, 235–237, 254, 317, 386, 402
  Raman optical activity (ROA), 4, 93, 114, 116, 217, 267
  Vibrational circular dichroism(VCD), 4, 84, 85, 93, 94, 105, 217
Statistical pattern recognition, 75, 287, 443
Surface-Enhanced Raman Spectroscopy (SERS), 4, 5, 29–51, 106, 138, 216, 254, 261–263, 284, 329, 402–415

T

Terpenes, 69, 256
Theoretical, 30, 38, 76, 91, 107, 125, 171–174, 274, 350, 436, 486
Theoretical calculations, 38, 68, 113, 193, 205
Tissue imaging, 8, 9, 291, 310
Toxicology, 387
Two-dimensional correlation analysis, 74, 216, 226, 237

U

Unicellular algae, 310, 313
Urea, 85, 220, 230–233
UV-Vis spectra, 94, 126

V
Vibrational Circular Dichroism (VCD), 3, 4, 84, 85, 84, 217
Vibrational spectroscopy, 3, 6–8, 108, 128, 215, 220, 281, 284, 355, 357, 364, 372, 373

W
Water, 5, 73, 84, 87, 105, 191, 216, 218, 223, 224, 289