Ion Beams in Nanoscience and Technology

Particle Acceleration and Detection

The series *Particle Acceleration and Detection* is devoted to monograph texts dealing with all aspects of particle acceleration and detection research and advanced teaching. The scope also includes topics such as beam physics and instrumentation as well as applications. Presentations should strongly emphasize the underlying physical and engineering sciences. Of particular interest are

- contributions which relate fundamental research to new applications beyond the immediate realm of the original field of research
- contributions which connect fundamental research in the aforementioned fields to fundamental research in related physical or engineering sciences
- concise accounts of newly emerging important topics that are embedded in a broader framework in order to provide quick but readable access of very new material to a larger audience

The books forming this collection will be of importance for graduate students and active researchers alike.

Series Editors:

Alexander Chao
SLAC
2575 Sand Hill Road
Menlo Park, CA 94025
USA

Christian W. Fabjan
CERN
PPE Division
1211 Genève 23
Switzerland

Frank Zimmermann
CERN
SL-Division
AP Group
1211 Genève 23
Switzerland

For further volumes:
http://www.springer.com/series/5267

Ragnar Hellborg · Harry J. Whitlow ·
Yanwen Zhang

Editors

Ion Beams in Nanoscience and Technology

With a Foreword by
Sir Michael W. Thompson

Springer

Editors

Ragnar Hellborg
Lund University
Department of Physics
Sölvegatan 14
SE-223 62 Lund
Sweden
ragnar.hellborg@nuclear.lu.se

Harry J. Whitlow
Department of Physics
University of Jyväskylä
FI-40014 Jyväskylä
Finland

Yanwen Zhang
Pacific Northwest National Laboratory
P.O.Box 999
Richland WA 99352
USA

Cover design: WMXDesign GmbH, Heidelberg

Printed on acid-free paper

To Gun, Ingegerd and Bill

Foreword

To make things, however large or small, one needs tools and measuring instruments. But before anything can be made by technology one must understand through science how the tools and instruments work. This book is for those who wish to make things that are so small that the position and size of individual atoms begin to matter, and where dimensions are measured in nanometres.

How can one control the making process for such things and how can we visualise what has been made? Such abilities came with the invention of microscopes based upon electron or ion optics and devices based upon ion scattering or sputtering that probe beneath a surface. These are instruments which can measure with nanometre precision.

What tools are available for the making process? Some came through understanding atomic collision processes under surfaces bombarded by energetic ions. Some others came from understanding how energy is lost from ions penetrating beneath the surface. That science gave us some of the tools for nanotechnology in the form of ion accelerators. With these one can precisely shape a surface by sputtering it away and modify the layer beneath by radiation damage. Also one can make new alloys or compounds in very thin layers by implanting energetic ions beneath the surface.

Such was the starting point for many endeavours reported in this book; and having been absent myself from any laboratory for many years, it is gratifying to see how much more of the science is well understood and how many things of value to the world can be made through this kind of technology.

From the beginning the electronics industry took a commendably long view by promoting this research in the universities as well as in its own laboratories. Exchanges of staff took place and collaboration ensured the opportunities for young scientists to pursue their careers through industrial research and development. Geographically the field initiated in North America, Europe, and Australasia but interest quickly spread to Asia, becoming a major activity in the burgeoning optoelectronics industries there. Industrial scale plants are commonplace in production units, and the newer applications extend into the biomedical sciences and electromechanical engineering. On the human side one could properly say there is now a third generation of scientists and engineers using ion beams in nanoscience and technology.

I expect that new understanding gained through current research will enable yet more new things to be made. This book should assist both activities. To those who will make the discoveries and inventions I express the hope that they may experience personal fulfilment through the realisation of their ideas.

Cornwall   Michael Thompson
June 2009

Editor's note: Professor Sir Michael W. Thompson was one of the early pioneers of atomic collisions in solids, which formed the seed from which the modern applications of ion beams in nanoscience and technology grew. His contributions include cornerstone developments in sputtering in amorphous materials and single crystals sputtering, surface scattering processes, and channelling. Subsequently, Sir Michael was appointed vice-chancellor of the University of East Anglia and then Birmingham before a period working in a leading position in the financial sector after retirement from academic life.

Preface

Energetic ion beam irradiation is the basis of a wide plethora of powerful research and fabrication techniques for materials characterisation and processing on a nanometre scale. Ion implantation is increasingly employed to tailor various material properties, focused ion beams can be used to machine away and deposit material on a scale of nanometres, and the scattering of energetic ions is a unique and quantitative tool for process development in high speed electronics. More recent developments are the use of MeV proton beams and other ions to fabricate 3D nanostructures with extreme aspect ratios for tissue engineering and nanofluidics lab-on-a-chip devices.

The potency of these energetic ion beam techniques lies in the electrostatic Coulomb interaction between energetic ions and individual atoms of the material. This allows substantial amounts of energy and momentum to be transferred to a single atom or electron in a way that is not possible with other probes such as energetic electrons, photons, etc. Since the start of the development of ion accelerators in the 1930s that were capable of producing high quality ion beams, energetic ion beam methods have played a leading part at the very forefront of the development of our modern high technology society. The technology behind these accelerators has developed in synergy with this technological development of materials. Dedicated ion implantation machines evolved from the early isotope separators by addition of sophisticated target handling systems. At the same time, spurred on by the quest for higher energy ions for fundamental nuclear science, higher energy machines evolved from the single-ended belt-driven machines to tandem accelerators. Since about the mid 1970s, more machines were installed for materials research than nuclear research. The requirements of energy stability and small machines led to introduction of the pellet-chain charged (Pelletron) and more recently rectifier-capacitor machines (first used by Cockroft and Walton in the beginning of the 1930s).

Certainly, the development of today's powerful personal computers and mobile communications would not have been possible without the use of energetic ion beams in nanotechnology. At the time of writing this book preface, society is facing new challenges, such as the development of sustainable technologies, pandemic disease control, and environmental issues. The broad range of applications and the usefulness of nanoscience and -technology methods using ion beams makes them really useful tools for meeting these challenges. Already they have been employed

for developing by mutagenisis high-yielding and resistant cash crops and developing efficient coatings for solar energy conversion.

In this work, established scientists from around the world have contributed deep knowledge in their respective fields. The goal is to produce a work that is not only a guide to practitioners, researchers, and students across the whole cross-disciplinary span of nanoscience and technology, but also serves to stimulate work in new areas where energetic ion beam methods can contribute to front-line research. Throughout the book the general recommendations of the International Union of Pure and Applied Physics (IUPAP) about units have been followed.

We are especially grateful to Professor Sir Michael Thompson for kindly accepting our invitation to write the Foreward to this book. Sir Michael was one of the early pioneers in the field of ion beams in nanoscience. Testimonials to his influence and relevance in the field are that his research on defects and radiation damage, understanding of the sputtering process, and also the scattering of ions from the surface layers in materials are reoccurring themes in many contributions to this book.

Lund, Sweden   Ragnar Hellborg
Jvyäskylä, Finland   Harry J. Whitlow
Richland, Washington   Yanwen Zhang
June 2009

Contents

Contributors

Eduardo Alves Department of Physics, Instituto Tecnológico e Nuclear (ITN), EN. 10, 2686-953-Sacavém, Portugal, ealves@itn.pt

Hans Henrik Andersen Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, DK-2100, Copenhagen Ø, Denmark, nimb@fys.ku.dk

K. Ansari Institute of Materials Research and Engineering, Singapore 117602; Singapore, mahabadik@imre.a-star.edu.sg

Kai Arstila K.U. Leuven, Instituut voor Kern- en Stralingsfysica, Celestijnenlaan 200D, Leuven B-3001, Belgium, kai.arstila@iki.fi

Andrew A. Bettiol Department of Physics, Centre for Ion Beam Applications (CIBA), National University of Singapore, Singapore 117542, Singapore, phybaa@nus.edu.sg

Daniel J. Blackwood Department of Materials Science and Engineering, National University of Singapore, Singapore, msedjb@nus.edu.sg

Mark Breese Department of Physics, National University of Singapore, Singapore 117542, Singapore, phymbhb@nus.edu.sg

Robert G. Elliman Electronic Materials Engineering Department, Research School of Physics and Engineering, The Australian National University, Canberra ACT 0200, Australia, rob.elliman@anu.edu.au.

Yongqi Fu School of Physical Electronics, University of Electronic Science and Technology of China, Chengdu 610054, P.R. China, yqfu@uestc.edu.cn

Fei Gao Pacific Northwest National Laboratory, Richland, WA 99352, USA, fei.gao@pnl.gov

Torgny Gustafsson Department of Physics and Astronomy, Rutgers University, Piscataway, NJ 08854, USA, gustaf@physics.rutgers.edu

Ragnar Hellborg Department of Physics, Lund University, Lund SE-223 62, Sweden, ragnar.hellborg@nuclear.lu.se

David Hole Department of Engineering and Design, University of Sussex, Falmer, Brighton, East Sussex BN1 9QH, UK, d.hole@sussex.ac.uk

Ee Jin Teo Department of Physics, National University of Singapore, Singapore, phytej@nus.edu.sg

Juhani Keinonen Department of Physics, University of Helsinki, Helsinki 00014, Finland, juhani.keinonen@helsinki.fi

Julia Laskin Pacific Northwest National Laboratory, Fundamental Science Directorate, Richland, WA 99352, USA, Julia.Laskin@pnl.gov

B. Gunnar Malm School of Information and Communication Technology (ICT), KTH – Royal Institute of Technology, Kista SE-16440, Sweden, gunta@kth.se

Sachiko T. Nakagawa Faculty of Science, Okayama University of Science, Okayama 700-0005, Japan, stnak@dap.ous.ac.jp

Kai Nordlund Department of Physics, University of Helsinki, Helsinki 00014, Finland, kai.nordlund@helsinki.fi

Antti Nuottajärvi Nanoscience Center, University of Jyväskylä, Jyväskylä FIN-40014, Finland, antti.nuottajarvi@jyu.fi

Mikael Östling School of Information and Communication Technology (ICT), KTH – Royal Institute of Technology, Kista SE-16440, Sweden, ostling@kth.se

Jyrki Räisänen Department of Physics, University of Helsinki, Helsinki FIN-00014, Finland, jyrki.raisanen@helsinki.fi

Timo Sajavaara Department of Physics, University of Jyväskylä, Jyväskylä FIN-40014, Finland, timo.sajavaara@jyu.fi

Tarmo Suppula Nanoscience Center, University of Jyväskylä, Jyväskylä FIN-40014, Finland, tarmo.suppula@jyu.fi

Peter Townsend Science and Technology, University of Sussex, Brighton BN1 9QH, UK, p.d.townsend@sussex.ac.uk

Christina Trautmann Gesellschaft für Schwerionenforschung (GSI), Darmstadt 64291, Germany, c.trautmann@gsi.de

Jeroen A. van Kan Department of Physics, Centre for Ion Beam Applications (CIBA), National University of Singapore, Singapore 117542, Singapore, phyjavk@nus.edu.sg

Peng Wang Pacific Northwest National Laboratory, Fundamental Science Directorate, Richland, WA 99352, USA, pengwang77@gmail.com

Lumin Wang Department of Nuclear Engineering & Radiological Sciences, Department of Materials Science & Engineering, Electron Microbeam Analysis Laboratory (EMAL), University of Michigan, Ann Arbor, MI 48109-2104, USA, lmwang@umich.edu

William J. Weber Pacific Northwest National Laboratory, Richland, WA 99352,
USA, Bill.Weber@pnl.gov

Tessica D.M. Weijers-Dall Electronic Materials Engineering Department,
Research School of Physics and Engineering, The Australian National University,
Canberra ACT 0200, Australia, tessica.dall@aihw.gov.au

Daniel J. White Max Planck Institute, CBG, Dresden 01307, Germany,
dan@chalkie.org.uk

Harry J. Whitlow Department of Physics, University of Jyväskylä, Jyväskylä
FIN-40014, Finland, harry.j.whitlow@jyu.fi

Yanwen Zhang Pacific Northwest National Laboratory, Richland, WA 99352,
USA, Yanwen.zhang@pnl.gov

Introduction

1 Energetic Ions and Nanometre Scale Interactions

What is nanoscience and technology? These words mean different things to different people. Often nanoscience is considered to deal with phenomena in the size interval between 0.1 and 100 nm. Technology is, in this context, the application of techniques over this size interval. One of the special characteristics of nanoscience are that there is a convergence between biomediciene, chemistry and physics. Although the smallest living organism is the cell, these are made up of biomolecules organised in a functional way. These molecules typically contain $100–10^5$ atoms and have a sizes of 1–30 nm. Virus particles haves sizes of 15–30 nm. In chemistry many molecules have sizes on a scale of nanometres. Also simpler molecules may be arranged in films and interface structures that have characteristic sizes on a nanometre scale. Knowledge of this had lead transistor development in electronics, following Moore's law [1], with gate lengths and gate oxide thickness that are on a scale of nanometres. From a physics viewpoint, nanometre structures are increasingly interesting because atomistic and collective quantum effects become important at this size level. The convergence of size scales where so-called "top-down" approaches such as the lithographic approaches that are used for nanoelectronics and "bottom-up" chemical and biological methods where the structures are assembled from smaller constituents, opens up a wide vistas for new research and development. Ion beam methods are very widely used across the whole nanoscience and technology field and have tremendous future potential because of their controllability and ability interact both collectively with many atoms or just a single atom in a selective way.

Energetic ions are charged atoms that have kinetic energy. They interact with the electrons and nuclei of material by Coulomb forces. The nature of the Coulomb interaction implies that energy from an energetic ion is transferred to a single atom, or electron. In other words, the primary interaction take place by excitation or scattering of individual atoms. Moreover, it is technically straightforward to control the energy, atomic species, direction and number of ions that impinge on a target surface. This degree of control, coupled with the well-established interaction probabilities between ions and atoms or electrons, *makes ions uniquely powerful tools*

for quantitatively and controllably analysing and modifying material over the entire field of nanoscience and technology. While up till now ion beam materials science and technology has developed as nanoscience and technology methods because of its applications in materials-chemistry, physics and electronics, the above mentioned convergence is finding increasing use in biomedical fields. It is interesting to note that biochemical and medical researchers today are main users of advanced physical methods such as protein structure determination by Nuclear Magnetic Resonance (NMR), synchrotron X-ray sources as well as neutron scattering.

In this book we have collected together a wide-range of the different techniques for materials analysis and modification on a nanometre scale. The emphasis has been on techniques, which provide well quantified results and techniques that are suited to working on the nanometre scale. In order to apply these ion beam techniques on a nanometre scale has spurred on development of instrumentation. Examples of this are magnetic lenses capable of focussing MeV ions to 10's of nanometre dimensions and spectrometers with advanced angular detection for correction of different geometrical effects.

2 Historical

The interaction of energetic ions with materials is usually associated with the development of radioactivity and nuclear physics. Even before the birth of nuclear physics the deposition of conductive coatings on the insides of discharge tubes was recognised in the mid 1850s [2] – although the process of sputter deposition of coatings from a plasma was not well understood till much later. Indeed since the sputtered atoms originate from the outer atomic layer or so, this can be considered to be one of the first observations of nanoscience, although it was not recognised as such until a century or so later. The next milestone were observations in the 1890s by Becquerel [3] and Mme. Curie that alpha particles were stopped by matter. Since the turn of the last century, the interaction of ions with matter was of key importance for nuclear science and this is evidenced by the list of top workers in the field including Bethe, Bohme, Bohr, Geiger, Marsden, Rutherford and Teller. The development by the initiation of a nuclear reaction using a beam from an accelerator (by Cockroft and Walton in 1932) was the next step taken forward. At about the same time, electrostatic accelerators that were developed by van de Graaff and Herb in 1930 were capable of producing ions with energies of MeV. A detailed historical overview of accelerator development is presented elsewhere by one of us [4]. Wartime and post-war research can be coarsely characterised in phases such as the atomic-era, electronic-era, main-frame and personal computer-eras and finally the nanoscience and technology era. Understanding of the problems of radiation damage in nuclear reactor and weapon materials was a prime concern in the atomic era. This necessitated study of the atomistic processes involved in defect formation and agglomeration by electron microscopy and other techniques constitutes one of

the foundations of modern nanotechnology. Isotope separators that were developed in the 1960s and 1970s provided a convenient source of energetic ions for these radiation damage studies and this lead to ion implantation. It was quickly realised that this provided a superior doping technology for the emerging semiconductor industry. This stimulated one of the most significant works in the field by Lindhard, Scharff and Schiøtt [5] who used a transport equation approach to determine ion implant range distributions in terms of dimensionless energies and depths. The transport equation approach was extended by Sigmund and co-workers to include the treatment of radiation damage, sputtering, multiple scattering and atomic relocation in mixing. The advantage of this analytic approach was that the necessary calculations could be made using a simple calculator.

In the 1970s and onwards to today, the needs of emerging the semiconductor industry stimulated microelectronic process development. The semiconductor industry used small electrostatic accelerators sputter profile methods such as Secondary Ion Mass Spectroscopy (SIMS) and Rutherford backscattering spectrometry (RBS) for a wide range of characterisation tasks such as measuring implantation profiles, diffusion, crystal defect annealing, oxide layer thickness and quality etc. The introduction of the mainframe computers was one of the results of the rise of the semiconductor industry, and one of the first uses of these was to numerically simulate ion implantation and radiation damage. The advent of numerical simulation overcame some of the difficulties of the analytic transport equation based methods such as handling multielemental inhomogeneous targets, which opened up new approaches to analysis of experimental spectra and calculating range and damage distributions using the binary collision approach. However, it was not till the 1990s when ion implantation and other developments, such as plasma etching, made the personal computer a reality. This advent of low-cost computing allowed widspread use of molecular dynamics simulations which offered detailed information about atomic defects and migration on a nanometre scale. The electronics revolution also gave sophisticated basic nuclear physics data collection systems. This made possible multidispersive Time of Flight Elastic Recoil Detection Analysis (ToF-ERDA), which combined with the improvement in numerical data handling, gave unambigious and clear data.

About the same time the term nanoscience starts to be widely accepted, which represents a convergence of physics, electronics, chemistry and biomedicine, started to become established. In this convergence is where phenomena take place on a scale that quantum and atomistic effects play a significant part in the properties of the system. In this period the development of DNA and protein sequencing lead to a major revolution in biomedicine. It became realistically possible to understand how genes are expressed and trace biological pathways. Early uses of ion implantation were to modify the corrosion properties of joint prosthesis by implantation and measurement of the mass of proteins by plasma desorption mass spectroscopy. The development is summarised in Fig. 1. Focused ion beams first became available in the 1980s and with Particle Induced X-ray Emission (PIXE) they have been used for trace element studies in biological matter at the tissue level which represents a size scale of 10 μm–1 mm. This has in turn evolved to Proton Beam Writing (PBW), a advanced nanometre lithography technique with a wide range of front-line applications.
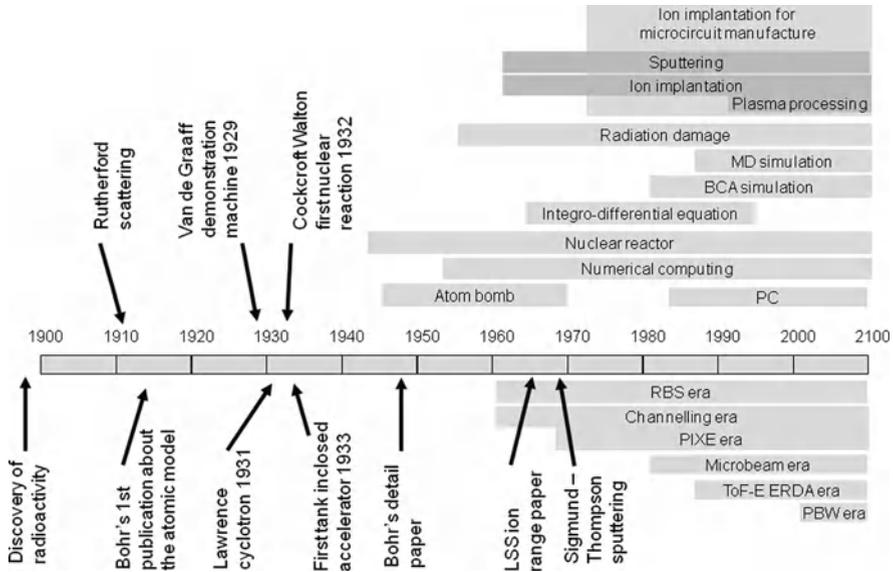
Fig. 1 Timeline illustration of the development of ion matter science and technology

3 Structure of This Work

The book is divided into five parts. These are comprised of chapters covering different areas and shorter box-chapters which focus on a specific theme. Part I consists of chapters that provide an overview of the current status of the fields and point out both the challenges to advancement in the field and also areas where ion beams can be applied in new ways.

Part II introduces the fundamental processes underlying ion-matter interactions in a series of chapters and boxes. The emphasis on modern approaches to calculating the parameters of interest by numerical computational models rather than conventional theory.

As nanoscience and technology based products appear on the marketplace, there will be an increasing requirement for quantitative analytical methods for development and quality control. Part III covers ion-beam based characterisation methods, including new methods for performing diagnostic tests on nanoclusters, which are important for optotronics.

The controllability of ion irradiation makes it a very attractive nanofabrication approach. The application range from individual sample processing in preparing samples for electron microscopy to mass production of optical and biomedical devices using stamp technology. This wide range of topics is presented in a series of boxes and chapters on different themes in Part IV.

The specialised equipment that is used for ion beam based research, development and production of nanometre scale materials is the subject Part V. In particular the

analytical techniques place stringent demands on instrumentation, especially if the ions are to be focused to nanometre dimensions. Furthermore, the electronics, (discussed above) has lead to new developments in electronics and control which are introduced in this part.

References

1. G.E. Moore, Electronics 38(8), (1965). Downloadable from: http://www.intel.com/museum/archives/history_docs/mooreslaw.htm
2. W.A. Grove, Phil. Mag. 5 (1853), 203.
3. A.H. Becquerel, Comptes Rendus 122, (1896), 420.
4. R. Hellborg, in: Electrostatic Accelerators (ed.) R. Hellborg (Springer, Berlin, 2005).
5. J. Lindhard, M. Scharff, H. Schiøtt, K. Dan. Vidensk. Selsk. Mat. Fys. Medd. 33(14), (1963).

Part I
Trends in Nanoscience and Technology

The objective of this part of the book is not simply to provide an overview of how ion beam techniques are used in nanoscience and technology but to give insight in new future possibilities for the application of ion beams in physics, chemistry and biomedicine. The large number of researchers who are active in these emerging field means that progress is rapid and it is difficult to come with hard and fast predictions. It is important to realise that the size scales involved in nanostructures are characterised by having sizes comparable with a few atoms to a few thousand atoms, which is the size of typical molecules representing the building blocks of matter. These building blocks can be assembled in different ways to realise tailored nanometre scaled devices – this is termed the *bottom-up approach* where chemical synthesis and biological gene expression are used to fabricate complex nanomachines such as zinc-finger proteins that regulate the production of expression of genes. The other approach, termed the *top-down approach*, is to carve-out and pattern nanostructures starting from bulk material. A good example is semiconductor process technology where complex circuits are mass produced. As such, the nanoscience and technology cover a convergence of the fields of biology, chemistry and physics. From these disciplines, sub-fields crossing traditional discipline borders are emerging. The authors in this section have given their own interpretation of present and future possibilities for application of ion beams in three main areas biomedicine, high-speed electronics and surface modification.

Nanoscale Engineering in the Biosciences

Daniel J. White and Harry J. Whitlow

1 Biological Scales and Biotechnology on a Nanoscale

1.1 Introduction

Biological matter is one of the most diverse and important classes of materials. Products of living organisms (wood, bone, cotton, wool, leather, coal, oil, drugs, etc.) are vital to humanity as foodstuffs, energy sources, engineering and construction materials, and chemicals; and by the way, they shape the environment of the biosphere.

Nanometre scale biology is concerned with cellular and subcellular levels. It is no understatement to say that rational modification of biomaterials on the nanoscale by methods in molecular biology and now by physical methods, gives us the potential to understand and control cells and entire organisms. The goals of this chapter are both to provide an introductory overview of nanobiology, and to illustrate how energetic ions are powerful tools used to modify and characterise biomaterials on this scale.

1.2 Molecular Biology Operates at the Nanometer Scale

All biological matter has the characteristic that its structure and distribution is encoded in genetic information stored in a base-4 code, carried in the sequence of four different nucleotide bases (A, T, G, and C) in different combinations of 3-letter words (codons) along polymer chains of deoxyribonucleic acid (DNA) [1]. There are 64 (4 cubed) different codons, redundant groups of which code for each of the 20 naturally occurring amino acids found in proteins. The sequence of the codons on the DNA strand of a single gene is translated into the sequence of the amino acid residues in the polypeptide polymer chain of a single protein. Tens of thousands of

D.J. White (✉)
Max Planck Institute of Molecular Cell Biology and Genetics (MPI–CBG),
Dresden 01307, Germany
e-mail: dan@chalkie.org.uk
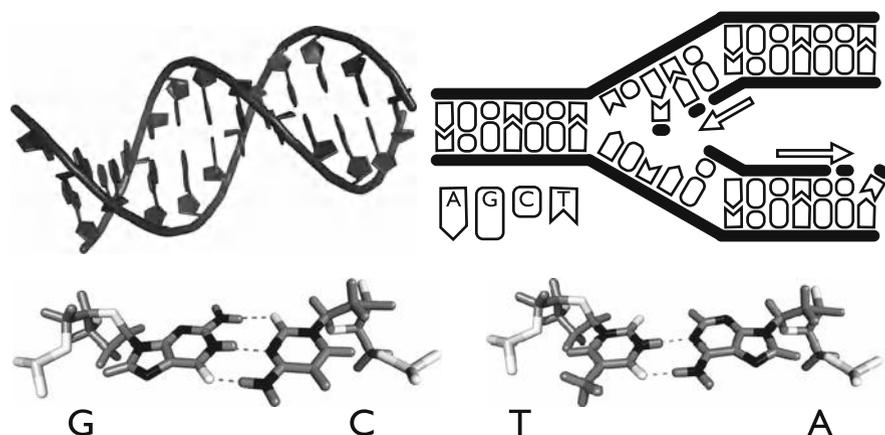
Fig. 1 *Top left*: the double helical structure of DNA. Sugar-phosphate backbones as tubes, with polygons showing the pentose sugar and heterocyclic base rings. *Top right*: DNA replication by synthesis of two new polynucleotide chains, using the original chains as structural templates. *Arrows* show the direction of DNA synthesis, opposite on either chain. *Bottom*: nucleotide residue structures with specific hydrogen bonding between heterocyclic bases on opposite polymer chains, always G≡C and A=T, reveal the elegant mechanism of sequence information preservation during DNA replication (atom shadings: N *black*, C and H *grey*, P *light grey*, O *white*). From this self-replication capability, DNA is a nanomachine in its own right

different proteins are encoded in the genes of a mammal's genome, but the simplest viruses have only 3 genes [2].

The B-form of deoxyribonucleic acid (DNA), shown in Fig. 1, has an antiparallel double helix structure $\sim$2.37 nm diameter where each turn, corresponding to a mean sequence length of 10.4 bp (base pairs), is $\sim$3.32 nm long. Single DNA molecules, as closed loops in some bacterial chromosomes, or in pairs of linear chromosomes in fungi, plants, and animals, may range in length from a few μm to many meters.

Proteins, along with ribose nucleic acids (RNA) in many critical functions, are the nanomachines of life, determining the form and function of all biomaterials. RNA structure is also encoded in genes. A typical protein molecule contains 50–2000 amino acid residues, with a size in the nanometre range, and a molecular weight of 5–200 kDa (1 Da = 1 atomic mass unit, u is $^1/_{12}$ the mass of a ^{12}C atom). The ribosome, a large RNA/protein nanomachine, which translates the genetic code, has a weight of many MDa, and is approximately 20 nm in diameter [3].

The information required for a polypeptide to fold into its unique three-dimensional structure as a mature protein is held at the codon sequence level in DNA, but is expressed in the protein's amino acid sequence itself (primary structure), such that hydrogen bonds form spontaneously within the protein, forming alpha helices and beta sheets (secondary structure) which themselves fold up into the tertiary structure of the protein by further hydrogen bonding, assisted by ionic and hydrophobic interactions, typically in an entropy driven process. Individual

proteins often form larger complexes with other proteins and/or themselves (quaternary structure), giving rise to multifunctional nanomachines such as enzyme (catalytic) complexes, signalling complexes, and large mechanical structures, e.g., collagen fibrils and the cytoskeleton [4, 5].

Protein structure is intimately linked to function, giving the exquisite specificity, selectivity, and sensitivity of biological materials. These biomolecular properties have been finely honed by the process of evolution by natural selection, whereby random mistakes (mutations) made during the replication of DNA are translated into changes in a protein's amino acid sequence. Mutations may lead to an enhanced protein structure or function (phenotype). The mutated gene, encoding that enhanced protein, gives the organism some advantage over its competitors which have the normal version of the gene, leading to its increased reproductive success. Over generations, the frequency of the enhanced gene will then increase in the population due to its selective advantage, meaning more and more organisms will have the enhanced protein working for them.

Ancient and widely accepted breeding methods result in artificially improved phenotypes generated by natural mutations, but without knowledge of the molecular genetic details, while relying on artificial genetic selection. Modern tools in biotechnology and molecular biology allow the rapid modification of DNA sequences, enabling knowledge-based engineering of protein structure and function for use in nanobiotechnology and other applications [6]. This is the basis for development of recombinant proteins and modern genetically modified organisms (GMO). Engineered DNA molecules can themselves be used as nanobiotechnology tools [7], and polypeptides can be chemically synthesised to produce novel small proteins [8, 9].

The functional characteristics of proteins rely on molecular recognition at the nanometer scale, Fig. 2. Enzymes are catalytic proteins which bind their substrates using spatially and chemically selective, high affinity bonding interactions. The unique three-dimensional structure of each enzyme allows it to recognise its target molecule (metal ion, small organic molecule, drug, DNA/RNA, or other protein) using a combination of a binding pocket/surface shape matching the target molecule (lock and key) and specific directional bonding interactions between the enzyme's amino acid residues and the target molecule. Amino acid side chains and the polypeptide backbone provide functional groups of atoms such as carboxylic acids, amines, amides, aromatics and aliphatics, which can donate numerous different bonds to the substrate, including ionic and hydrophobic interactions, and hydrogen bonds [10]. The same is true not only for enzymes, but for all protein interactions with other biological macromolecules such as DNA/RNA, complex carbohydrates, and lipids. Thousands of different protein–protein interactions critical in cells are governed by these rules, where a closely fitting shape, often burying a large proportion of the surface area, and numerous specific bonding interactions are key [11].

Living organisms in their simplest form are single bacterial cells, such as an *Escherichia coli* bacterium. Proteins and other biomolecules act alone or form complex nanomachines with specific biological functions within the cell or incorporated into the lipid membranes surrounding the cell. Bacteria are prokaryotes, meaning
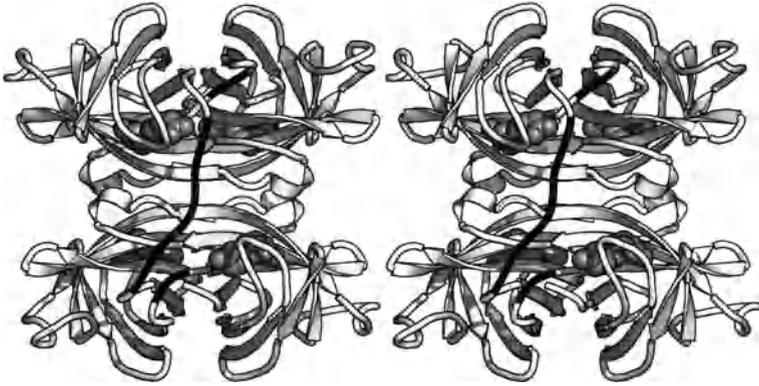
Fig. 2 Single chain avidin, a tool for molecular biology and nanotechnology, is a genetically engineered protein derived from chicken white avidin, which binds the vitamin biotin with the highest affinity constant known for a biological interaction. Biotin can be covalently linked to other molecules, acting as an affinity tag. Natural avidin is a tetramer, but single chain avidin is a single engineered polypeptide containing four avidin units with four biotin binding sites, which can now be independently mutated to change their ligand binding specificities from biotin to other small molecule tags. The protein is shown in a stereo ribbons representation following the alpha carbon trace of the polypeptide chain, with the three engineered linkers in *black*, and with the four biotin ligands in *grey* spacefill representation

their cell contains no nulceus. More complex organisms such as a sunflower plant, *Helianthus annuus*, or a human, *Homo sapiens*, are eukaryotes, meaning their cells contain a nucleus and organelles with internal lipid membranes. Eukaryotes may be single-celled, such as brewers yeast *Saccharomyces cerevisiae*, or complex multi-cellular organisms, with different types of cells organised hierarchically into tissues and organs with specific biological functions, for instance, muscle for movement or leaves for photosynthesis. An ensemble of organs makes up a complex organism such as a mammal. These hierarchical units span a range of physical sizes, as shown in Fig. 3.

In short, the structure and function of all living organisms from sub-μm sized bacteria to massive trees and mammals, and their products, have the common reliance on genetic information coded onto nanometre scale units of DNA (or in the case of some viruses, RNA), which is expressed in the functionality of nanometer scale proteins.

1.3 Cell Structure and Compartmentalisation of Function in Organelles

Cells are enclosed in a lipid bilayer membrane, the plasma membrane or cytoplasmic membrane. In bacteria, the cell is contained inside two membranes, with the periplasmic space in between these inner and outer membranes. Fungi, plant, and animal cells are composed of a wide variety of different nanomachines (organelles)
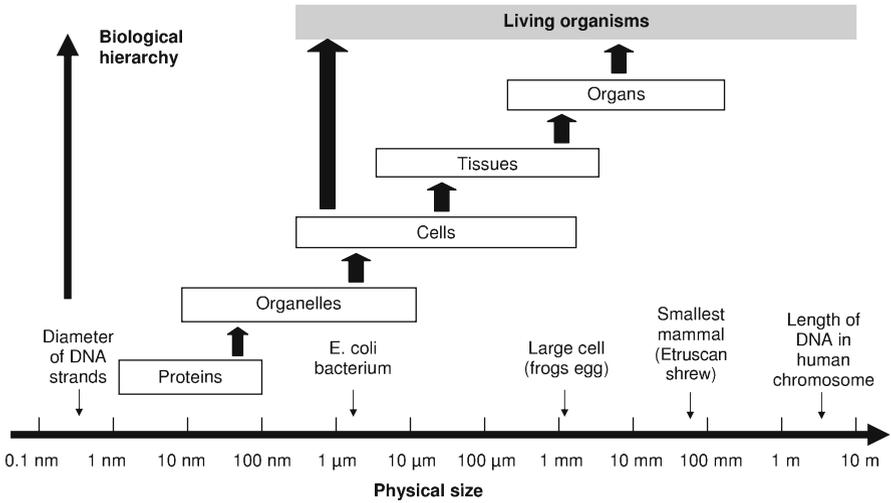
Fig. 3 Schematic illustration of biological scales

organised to compartmentalise specific biological functions, each often inside a lipid membrane.

Organelles are ensembles of proteins, lipid bilayer membranes, and in some cases RNA. For instance, the double lipid membrane-enclosed nucleus carries the cell's chromatin/chromosomes (DNA-protein complexes) and is the site of gene expression (transcription of a gene's DNA sequence into a complementary sequence of messenger RNA, mRNA). The centrosome is concerned with the segregation of chromosomes into the daughter cells during cell division. Ribosomes, often found attached to the membrane structures of the endoplasmic reticulum, are the protein factories of the cell. These RNA-protein nanomachines translate genetic information, carried from the nucleus to them in mRNA, into protein structure [3]. Many cell compartments, such as the endoplasmic reticulum (ER), peroxisomes, lysosomes, and transport vesicles, are membrane-bound microenvironments containing specific enzymes and chemical properties such as pH, redox level, etc. Mitochondria are double lipid membrane-enclosed organelles containing protein nanomachines responsible for the transfer of chemical energy from the oxidation of sugars into adenosine triphosphate (ATP), the energy currency of the cell. Mitochondria and plant chloroplasts are considered to be evolved from endosymbiotic bacteria, since they have their own separately replicating genetic information in a mini-chromosome encoding many of the proteins that they contain, and indeed, physically and functionally resemble existing types of bacteria. Figure 4 shows a schematic illustration of a generic animal cell.

Bacteria do not have membrane-bound organelles, but carry their chromosomes in the cytoplasm. Plant cells share nuclei with animal cells, but are enclosed by a rigid cellulose cell wall, allowing osmotic pressure from their large fluid-filled vacuole to make the cells turgid. Plant cells carry photosynthetic chloroplasts, which
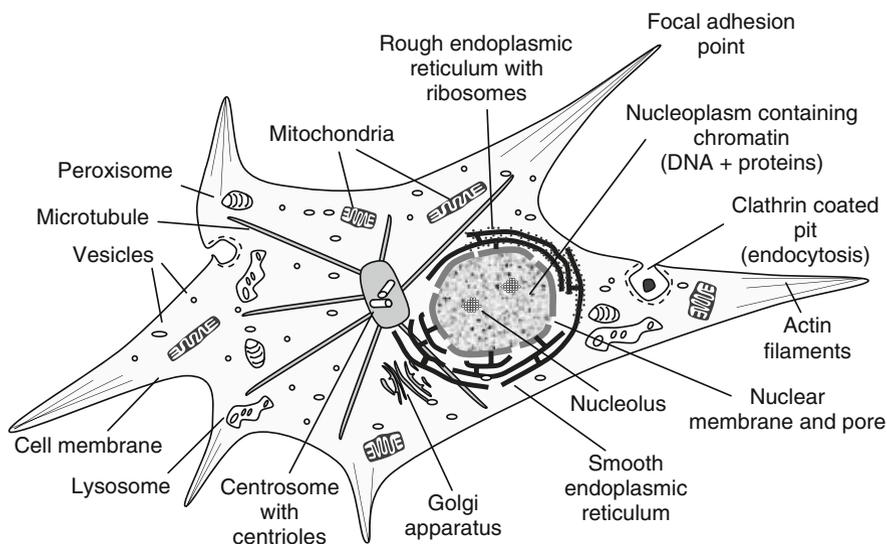
Fig. 4 Schematic overview of a generic animal cell illustrating different organelles

carry out the conversion of light energy into chemical energy stored in sugars. Fungi cells are similar to plant cells but lack chloroplasts and large vacuoles.

Viruses, while not being technically alive in that they are not free-living or able to replicate without the assistance of a host cell, do carry their own genetic information in DNA or RNA. In fact, viruses are incredible nanomachines, capable of infecting a host cell by forcing or tricking the cell to let their DNA/RNA enter for replication by the host cell's own machinery. New viral particles, virions, are self-assembled by packaging the nascent viral DNA/RNA into the viral coat proteins synthesised by the host's ribosomes using the viral genetic information. Many viruses also contain lipid membranes stolen from the host cell.

2 Modern Tools in Nanobiotechnology

2.1 Molecular Biology Methods in Nanoscience

Over the past few decades, rapid developments in biotechnology have allowed rapid protein and DNA structure engineering. Central to this progress have been advances in bacterial genetic engineering which allow isolation or "cloning" of different genes, and, importantly, the polymerase chain reaction (PCR) which allows the *in vitro* amplification of specific genes or other DNA sequences, and also mutagenesis of the gene's DNA sequence and resulting protein structure.

Engineered DNA molecules can be produced on a large scale by PCR or chemical synthesis, followed by purification using cheap gel electrophoresis or liquid phase

chromatography methods. DNA molecules which self-assemble into reproducible nanostructures of a designed shape [7], by virtue of the sequence specific binding of complementary strands, have been characterised recently, opening the door for their use in nanofabrication when chemically modified to carry functional groups which bind specifically to other nanoparticles and substrates.

The PCR reaction uses thermostable DNA polymerase enzymes from thermophillic bacteria such as *Thermus aquaticus* or *Pyrococcus furiosus* along with a pair of specific oligonucleotides (short single-stranded DNA molecules) which anneal to the heat denatured template DNA strands in a sequence specific manner, and prime DNA replication from their ends in a thermocyclic chain reaction. It use of oligonucleotides with sequences complementary to mutagenesis target sequences, but carrying a small number of nucleotide substitutions, insertions, or deletions, allow a PCR reaction to amplify genes in vitro incorporating mutated sequences. Proteins can now be redesigned, using *in silico* molecular modeling and mechanics methods, facilitated by the increasing power of modern computers, to predict their enhanced or new functions. It is possible to redesign a protein's binding specificity for small ligands and other macromolecules or nanostructures for use in nanotech applications. However, even relatively small proteins can not be designed from scratch. Rather, new functionality must be designed into protein structure by modification of an existing protein. Present computing power and available molecular dynamics methods are insufficient to compute the correct 3D folded structure of a protein starting from an extended polypeptide chain (the "Levinthal Paradox").

When engineered genes are decoded or "expressed" in cells, mutated proteins are produced with the desired altered structures. Various protein expression systems are available using bacteria, yeast, insect, and mammal host cells. Bacterial proteins are easiest to produce in large quantities in bacterial cells. Mammalian proteins are often produced in a misfolded insoluble form in bacteria as a result of the lack of mammalian protein folding "chaperone" proteins and protein modification enzymes (e.g., glycosylation and acylation) in bacteria. However, mammalian proteins can be produced in native forms in insect and mammalian cells, and sometimes more simply in yeasts. Insect cells are easy to grow in suspension culture and can express "heterologous" mammalian proteins with very similar structure and function to the native protein from a mammalian cell, having been modified similarly by, for instance, glycosylation. Different small artificial chromosomes (plasmids) for each type of protein expression system must be constructed by DNA engineering, each carrying growth selection genes (e.g., antibiotic resistance genes) for maintenance in bacteria and protein expression host cells, and the gene encoding the protein to be produced. Regulatory DNA sequences must be included, allowing the metabolically triggered expression of the genes and replication of the plasmid in bacteria during engineering and possibly also in the host cells.

Proteins expressed at lab scale (up to 10s of litres) at high levels in cells can be purified to homogeneity in mg quantities for use in downsteam applications. This process is often aided by removable affinity tag proteins, genetically fused to the protein of interest, for ease of purification (e.g., His_6 tags which bind to Ni^{2+} chelation resins). Inexpensive one-step liquid phase chromatography purification of

recombinant protein from cell extracts is now commonplace. Recently, green fluorescent protein (GFP) has been increasingly widely used as a fluorescent fusion tag on proteins of interest. GFP-tagged proteins usually retain their structure and function and behave normally in cells, where they can then be imaged by fluorescence microscopy, enabling the distribution of the protein under study to be visualised in living cells and other microstructures. Capsid (viral coat) proteins can be expressed in insect cells and purified as functional virus-like nanoparticles [12]. Purification of genetically engineered virus-like particles containing cargo other than the native DNA or RNA, such as quantum dots or drugs, is now possible due to their self-assembly and molecular recognition properties. These biological nanoparticles can be made to adhere to specific biomolecules in target tissues (e.g., cancer cells) or nanofabricated materials, by genetically engineering their capsid surfaces, altering their binding specificities.

2.2 Spectroscopic Methods in Biomolecular Nanoscience

The atomic details of biomolecular structure can be revealed by a variety of spectroscopic methods, including X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy. Crystals can be grown from solutions of purified protein and small viruses, which diffract X-rays, allowing determination of their 3D atomic structure. Purified proteins produced by organisms growing on nutrients containing ^{15}N and ^{13}C can be analysed by multidimensional NMR, again allowing 3D atomic structure determination, in this case in an aqueous solution state, with added dynamic and chemical information, without crystal packing artifacts. Electron paramagnetic resonance (EPR) spectroscopy can be used to probe the electronic properties of paramagentic metal ions coordinated in catalytic or structural metal binding sites in proteins, giving information on how protein coordination modulates the electronic structure and chemistry of the metal ion.

   Modern mass spectrometry methods, such as Matrix-assisted laser desorption/ ionisation – Time of Flight (MALDI-TOF), are invaluable tools in biomolecular studies. Isothermal titration calorimetry (ITC) is a robust method for determining the thermodynamic parameters for reversible ligand binding to proteins in solution, allowing direct access to the thermodynamic dissociation constant (K_d). Indirect methods, such as fluorescence and circular dichroism titrations, can be used to measure K_d, but are prone to uncertainties. Surface plasmon detection methods (such as Biacore and IAsys), enable the rapid measurement of protein-ligand $K_{off/on}$ values, in a framework allowing high throughput screening. Differential scanning calorimetry (DSC) probes the thermodynamic properties of proteins with respect to their folding energy profiles and thermal stabilities, which can be altered by ligand binding and protein–protein interactions. Fluorescence correlation spectroscopy (FCS) is a single fluorescent particle detection method, that measures the diffusion coefficient of a protein or small molecule, providing an estimation of its molecular size. FCS is useful in ligand binding and complex formation measurements with fluorescence labeled molecules, where the change in fluorescent particle size after binding can

be measured. Atomic force microscopy (AFM) is another single molecule method. In force spectroscopy mode one can measure, under physiological conditions, the dynamic interaction forces between purified biomolecules and their ligands, the stretching and unfolding forces of DNA and proteins, and also the forces involved in the adhesion of living cells to immobilised receptor molecules, all in the pN range.

2.3 Imaging Methods in Biomolecular Nanoscience

In biological imaging, negative stain electron microscopy (EM) has been the mainstay of nanostructure analysis, revealing the structure of cellular organelles. Target molecules can be labeled with gold nanoparticles attached to target specific antibodies, allowing the target molecule's distribution in cells to be visualised. However, artifacts from sample fixation and negative staining can be a problem, especially for small, delicate protein/membrane structures. Recent advances in fresh sample high pressure vitreous freezing and low sample temperature EM technology (cryoEM) allow imaging of well preserved structures, such as whole viruses, leading to 3D virus structure determination to low nm resolution [13] and tilt-series 3D reconstructions of subcellular objects [14].

Optical microscopy rules the μm domain, with cell biology now relying heavily on fluorescence microscopies, especially confocal microscopy for small structures down to the diffraction limit, to gain structural and localisation information on biomolecules, using target specific fluorescent antibody labeling and fluorescent protein fusion tags e.g., GFP. Modern confocal microscopes can take many frames per second during live-cell imaging, and can image several different fluorophores simultaneously using spectral imaging or different excitation laser lines and emission filter/detector channels. Methods such as fluorescence recovery after photobleaching (FRAP), fluorescence loss in photobleaching (FLIP), fluorescent resonant energy transfer (FRET) and fluorescence lifetime imaging microscopy (FLIM) [15] give dynamic information on the movement of fluorophores and the proximity of different biomolecules to each other, both key questions in cell biology. AFM is being used to image, measure forces in and manipulate biomolecules such as DNA and collagen [16], viruses, and living cells, and can be combined with an inverted confocal microscope. Scanning near-field optical microscopy (SNOM) promises sub-diffraction limit fluorescence microscopy of living cell membranes at nanometer resolution. Recent developments in fluorescence microscopy such as stochastic optical reconstruction microscopy (STORM), photoactivated localisation microscopy (PALM), and stimulated emission depletion microscopy (STED) also promise improved spatial resolution.

Ion beams offer interesting and hitherto unexploited possibilities for biomedical imaging. MeV ions have straight trajectories in up to $10\,\mu$m or so of biological materials and deposit most of their energy within a few nanometres of the ion track. The short de Broglie wavelength of ions means that the diffraction which limits the attainable resolution in optical microscopy is negligible, and the attainable resolution is essentially limited only by the size of the spot (10–20 nm) that the beam

Fig. 5 Direct-STIM image of a fixed and critical-point dried cultured MCF-7 breast cancer cell (American Type Culture Collection [ATCC]). The grey scale has been expanded to emphasise the internal structural details of the cell., M. Ren, H.J. Whitlow, J.A. van Kan, A.A. Bettiol, D. Lim, Y.G. Chan, B.H. Bay, T. Osipowicz and F. Watt. unpublished data
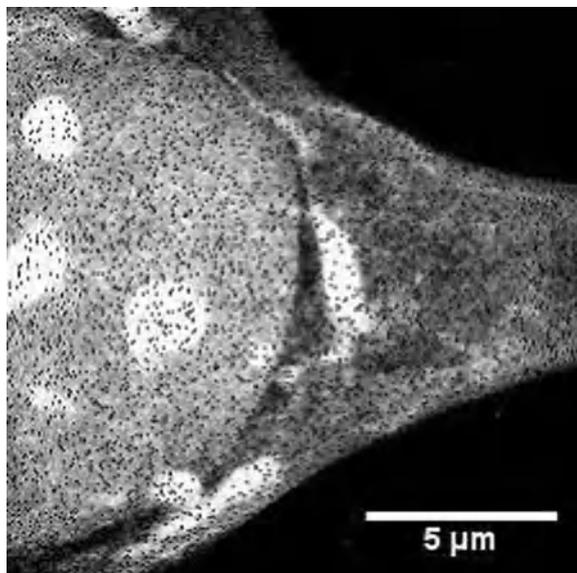
can be focused to. Using focused MeV ions for imaging cells with resolutions on a scale of nanometres requires significant developments beyond conventional particle induced X-Ray emission (PIXE) and particle induced gamma emission (PIGE) mapping. This is because as the beam is focused to smaller spot radii, r. The fluence and thereby also the dose increases as $1/r^2$. This, combined with the extremely low quantum efficiency of the measurement (many orders of magnitude less than one detected photon per incident MeV ion), limits the practical resolution attainable in PIXE and PIGE mapping to about the same as an optical microscope. Using high quantum efficiency signals for measurement, such as direct-scanning transmission ion microscopy (STIM) [17, 18], allows the fluence required for imaging to be reduced, and hence a smaller spot size can be used. This is illustrated in Fig. 5, which shows the structural information in a direct-STIM image of a cultured human breast cancer cell. The resolution is ∼50 nm.

3 Ion Bombardment of Biological Materials

One of the largest applications of ion irradiation in biomedicine is for cancer therapy, with usage of proton and He beams of 230 MeV per nucleon, and more recently 320–430 MeV per nucleon ^{12}C, as an alternative to X-ray therapy. This is outside the scope of this chapter; however a good overview is given in an IAEA report [19].

Biological materials are inherently labile and typically contain large quantities (70–95%) of water, which is essential for organism function. Ion accelerators and irradiation generally takes place under high vacuum ($10^{-1} - 10^{-6}$ Pa). Ion irradiation of plant cells is possible because their cellulose cell wall enables them to

survive a short period in vacuum [1]. The morphology of biological material such as cells is also modified by ion irradiation, forming small pores in surfaces such as skins and cell walls, helping foreign molecules enter the cells, which affects their viability. Some cells, such as animal and bacteria cells which do not have a cellulose cell wall, are much more susceptible to dehydration in vacuum [1]. This limits the possibilities for *in vivo* irradiation to sufficiently high energies (MeV) that the ion beam can be introduced into a cell growth environment through a thin window. The environmental chamber maintains the cells at almost natural conditions during the irradiation. In order to avoid stressing the cells, it is necessary to maintain them at very closely controlled temperature, humidity, and CO_2 range. In cell culture, the environment is typically maintained at 70–90 % relative humidity, $37 \pm 0.1°C$, in sterile air with a 5% excess of CO_2. During the short time of irradiation the residual cell growth media in the petri dish should be sufficient to provide the cells with nutrients and growth factors. This technique has found much application in single cell irradiation studies, e.g., [20, 21]. Irradiation using vertical ion beams, such as at the Gray Cancer Institute, allows the cells to grow normally in petri dishes with thin bottom surfaces.

For blanket irradiations of attached cells, where motion of the cells is of no importance, horizontal ion beams can be used. In this case the cell growth medium is decanted away and the irradiation performed quickly in air, then the cell growth medium is replaced immediately afterwards. Figure 6 illustrates schematically a beam line set up used in Jyväskylä for horizontal irradiation of mammalian cells. To ensure homogeneous irradiation the beam is rastered over the sample. Magnetic scanning is preferable, especially for energetic heavy ion irradiation where the rigidity of the beam is high. Determination of the fluence (and hence the dose) in the beam diagnostics section is of critical importance. The absolute measurement of the beam current can be done destructively using a removable Faraday cup with proper secondary electron suppression that intercepts the entire irradiated area. During irradiation, the use of scintillation detectors that intercept a fraction beam outside the irradiated area can be used to indirectly monitor the beam flux. The beam enters the irradiation chamber through a window that separates the high vacuum in the accelerator beam line from the atmosphere in the environmental chamber. Suitable foil materials are mylar or havar (high strength nonmagnetic Co-Cr-Ni-Fe-W-Mo-Mn alloy). In the event of a rupture of this window, a catastrophic ventilation of the accelerator will occur. In order to prevent this, a vacuum sensor is placed close to the window. This is connected to a distant fast-acting gate valve, which can shut before gas sufficient to cause damage escapes.

4 Ion Induced DNA Modification and Mutation

Ions interact with matter by depositing energy in electron excitation and physical displacement of atoms, as discussed in chapter "Basis of Ion Scattering in Nanoscale Materials". Ions deposit energy into the immediate locality of single molecules.
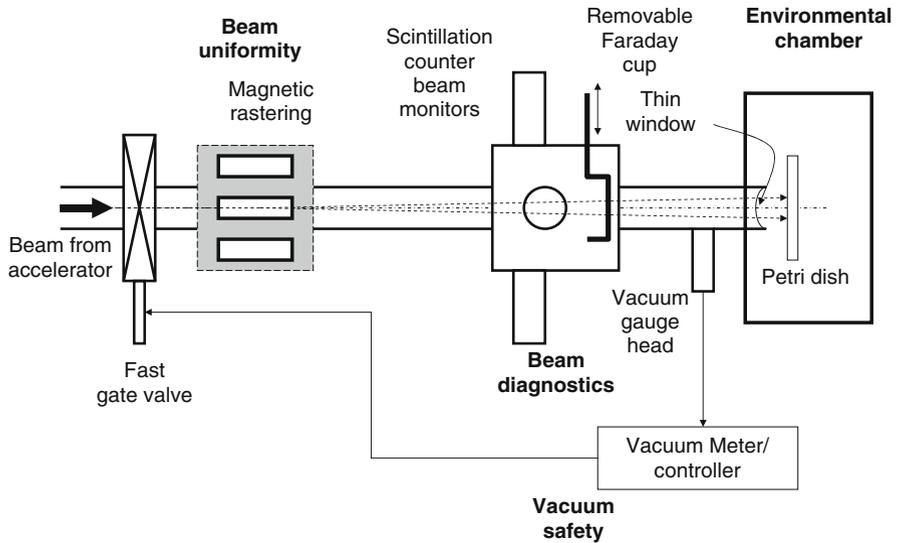
Fig. 6 Schematic illustration of a irradiation facility for horizontal ion beam

If this occurs at the chromosomes, it results in induced modification of the DNA sequence, altering gene expression and/or protein structure in the cell. This may be so damaging that the cell dies. Normally, natural DNA repair mechanisms correct most of this damage. However, the residual damage will give rise to permanent modifications in the DNA sequence or chromosome structure, known as *mutations* [1].

Cells have a high water content (70–95%). Electron excitation and atomic displacement both dissociate water, forming $OH^.$ radicals, which can undergo reactions with nucleotides within the aqueous environment of a cell, releasing and damaging nucleotide bases. $OH^.$ radicals can dehydrogenate deoxypentose in DNA, causing a break in one of the backbone strands (SSB – single strand break). SSBs can also arise in the DNA repair process after repair enzymes have converted damaged bases to apyrimidine/apurine (AP) sites, triggering DNA polymerase β (pol β) to cut the phosphate backbone. Radiation can also alter the DNA nulcelotide bases. For example, UV light can induce two adjacent thymines to form a thymine dimer, which causes an error in DNA replication [1]. However, the greatest effect of ion irradiation is double backbone breaks (DSB – double strand breaks), breaking the chromosome in two. Broken chromosomes and incorrectly repaired rejoined broken ends cause major genetic rearrangements and defects. One of the difficulties in using low-energy ion beams to induce changes in organisms is that the beam has a short range, generally requiring irradiation to be carried out in vacuum [1]. This limits the irradiation to organisms that have a high viability after exposure to vacuum, e.g., plant seeds and many spores. Irradiation of seeds and spores has been demonstrated to be a very successful way of introducing mutations into cereal crops and bacteria for biotechnological production (e.g., functional foods, pharmaceuticals,

Fig. 7 Illustration of rice stalk length genetically modified by nitrogen irradiation of seeds. Three months after germination. From Prof. Thiraphat Viliaithong, used with permission

and biochemicals). Figure 7 illustrates rice plants that have undergone genetic modification by irradiating seeds with low energy (30 keV) nitrogen ions.

Tuberculosis, a global health epidemic issue, can be treated with rifamycin, synthesised by the bacterium *Amycolatopsis mediterranei*. The rifamycin production titer could be increased from 20% to 90% by ion implantation over multiple generations of the bacterial spores [1]. Genetic cloning requires the transfer of engineered DNA into target cells. Conventional physical and chemical methods to do this are the gene gun (bombardment cells with DNA-coated micro- and nanoparticles), electroporation (a high transient electric field creates perforations in the cell wall and membrane, allowing entry of biomolecules), calcium chloride shock-induced DNA uptake, and also liposome transfection reagents. These methods are toxic to the cell, resulting in cell death. Further, transfer or transfection efficiency is low. Ion beams increase the permeability of the cell wall/membrane of plant and bacterial cells, and also activate DNA repair mechanisms because of ion induced chomosomal damage, increasing the efficiency of stable gene transfer, where genetic changes are passed on to subsequent generations. A good example of this is the transfer of the green fluorescent protein (GFP) gene into bacteria (*E. coli*) [22]. Fluences up to 2×10^{15} 26 keV nitrogen or argon ions cm^{-2} allowed transformation with several 3 kbp plasmid DNA constructs.

5 Ion Induced Modification of the Cell Environment

Tissue is built up in a genetically organised manner within the extra-cellular matrix (ECM) by cell-cell and cell-ECM interactions, which may be highly specific or nonspecific. Specific interactions include collagens and fibronectin with integrins, which are involved in cell signalling, enabling the cell to sense the composition of the ECM. In this way stem-cell differentiation is partially controlled by ECM composition.

Ion bombardment deposits large amounts of energy, leading to atomic displacement and thus chemical alteration of specific cell receptor (e.g., integrin) binding sites on the ECM molecules (e.g., collagen) which could disrupt receptor binding and cell signaling pathways [23]. This could be used to engineer cell adhesion and proliferation.

Nonspecific adhesion of cell surfaces is mediated by numerous cellular receptors on the cell membrane, which nonspecifically bind glycoproteins excreted onto the surface by cells, thus forming their own nonspecific ECM. Ion beam deposition of, e.g., diamond-like carbon (DLC) or modification of the surfaces of polymers by ion beams, creates surfaces which may positively or negatively affect cell adhesion and proliferation, allowing cell patterning. A good example is neuron proliferation on lithographically-patterned DLC/glass substrates [24]. Figure 8 shows the proliferation of cultured neuron cells on DLC coated and uncoated regions of a glass slide. DLC films are extremely hydrophobic, which stimulates attachment through promoting excretion of ECM.

Typically cells are cultured on flat surfaces of glass or plastic. However, natural environments are 3-dimensional, with features on the scale of nm to mm. Thus, growing cells in engineered 3D environments should provide a more natural environment than an alarmingly unnatural flat surface. Cells move by attaching to specific ECM ligands at the leading cell protrusions and then contracting the cytoskeleton to bring the trailing edge forward [25]. A textured 3D surface can be created by ion bombardment and MeV ion beam lithography (chapter "Proton Beam Writing: A New 3D Nano Lithographic Technique", Box 11). Such surfaces provide a greater number of potential adhesion sites, even before coating with specific ECM ligands. As a result, cells will tend grow into and move over such surfaces in preference to flat surfaces. Ion beams can be used to fabricate such 3D structures. Figure 9 illustrates a 3D nanostructured surface produced using ion irradiation [26]. On surfaces with ion beam fabricated 3D microchannels, the cytoskeleton of mouse pre-osteoblasts is organised along the surface ridges [27].

Fig. 8 Preferential spreading of neurons on DLC on a glass slide. The *upper part* above the *dotted lines* is coated with a DLC film, *below* is uncoated. After Brown et al. with the permission of Institute of Physics Publishing [24]
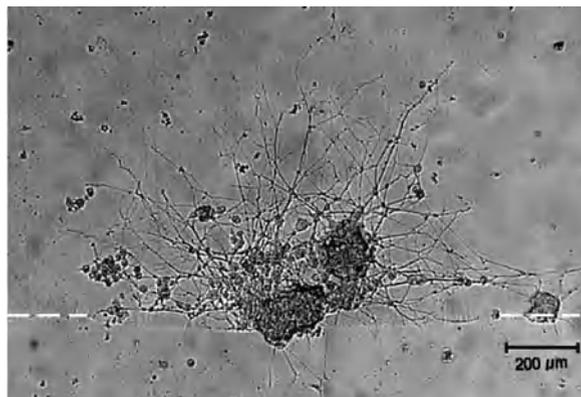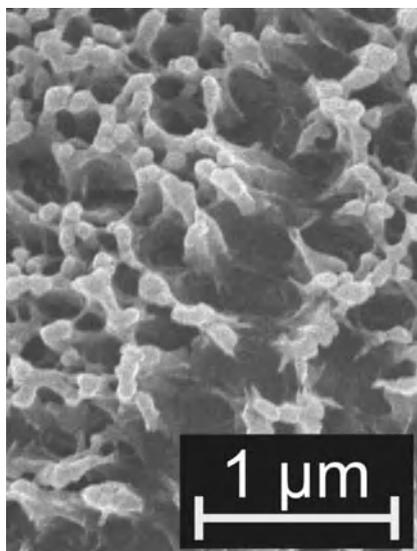
Fig. 9 Example of a
nanostructured PMMA
surface produced by 56 MeV
N^{3+} irradiation followed by
development in 7:3
isopropanol-water mixture.
After Gorelick et al. [26]

6 The Future of Ion Beams in Nanobiomedical Sensors, Labs-on-a-Chip, and Tissue Engineering

Miniaturisation opens many possibilities in biomedical technology. Examples of such systems are lab-on-a-chip bioassay kits and microspotting plates. A lab-on-a-chip system is comprised of devices such as microfluidic channels, pumps, valves, reservoirs, biosensors, optics, and also micro-electro-mechanical systems (MEMS) devices. For many applications, low manufacturing costs, uniformity, low false positive results, sensitivity, reproducibility, speed, and simplicity are all key considerations. In particular, the sensitivity and specificity must be extremely high in order to detect small numbers of specific biomolecules, for example, viruses. In order to achieve this, biological sensors such as antibodies are often used.

MeV ion beam lithography (chapter "Nanoscale Materials Defect Characterisation) has advantages over conventional methods for fabricating such miniaturised analysis systems because it can be used to pattern 3D structures in thick resist with a high aspect ratio. The large penetration implies channels can be directly written with sufficient depth to allow the passage of, e.g., biomolecules, viruses, and even whole cells with sizes indicated in Fig. 3. A further key advantage is that the lithographic technique can be used to produce stamps for mass production by, e.g., hot embossing. The biosensor is the heart of the system. This can be based on a number of physical phenomena such as electrical impedance, as well as optical, magnetic and surface plasmon properties. These rely on chemical functionalisation of a surface such that the biological sensor molecule (antibody) can be attached and immobilised on the surface. A more simple device is based on the ELISA (enzyme linked immunosorbent assay) technique, where detection of a biomolecule by enzyme linked antibody catalyses a colour change reaction. This could be

combined with MeV ion beam lithography produced microfluidics to increase sensitivity and/or sophisticated pretreatment (e.g., to separate red and white blood cells, or blood cells from plasma, or even chemical pretreatment such as pH and salinity adjustment).

Fluid transport is an important aspect in microfluidics design. Fluid pumping can be based on capillary forces, electokinetic methods [28], pressure driven method, thermal expansion of a reservoir, or MEMS pumps. Use of the 3D machining properties of MeV ion beam lithography make fabrication of complex 3D microfluidic structures with buried reservoir and vias possible. The difficulty with this is that mass production by stamping is not feasible. However, this restriction can be to some extent mitigated by sealing the top of a channel with an overlying film after hot embossing. Valves can be realised in various ways, for instance, single-operation valves based on melting wax plugs by electrical heating, repetitive valves using pressure-actuated membrane valves, or passive valves based on hydrophobic constrictions [29].

Another area is the engineering of surfaces to improve corrosion resistance in medical prostheses such as hip joints, heart valves [30], and artificial organs. Body fluids are very corrosive and can attack even surgical grade steels, leading to immune rejection. Ion implantation, e.g., with nitrogen or carbon ions, can modify the surface, which not only improves wear resistance, but can also modify the surface passivation, i.e., increasing or decreasing immunological tolerance [31, 32]. In this connection, nanostructuring a surface with ion beams may assist tissue growth and adhesion *via* stimulation of fresh ECM deposition.

7 Concluding Remarks

Ion beam methods provide a wide range of different possibilities for biotechnology and medicine. The complexity and labile nature of biological materials requires novel thinking and interaction between ion beam scientists, biologists, and nanoscientists with a cross disciplinary approach. Areas where development appears particularly fruitful are:

– Ion induced gene transfer and mutation of eukaryotic or prokaryotic cells and organisms for gene therapy and biotechnology.
– Economical protyping and mass production of microfluidic devices.
– Nanoscale bioimaging with ions at high resolution, filling the resolution range gap between electron and light microscopies.
– Surface modification for engineering biosurfaces.

References

1. Z. Yu. Introduction to Ion Beam Biotechnology. Trans. L. Yu, T. Vilaithong, I. Brown (Springer, New York, 2006).
2. B. Lewin. Genes VIII (Prentice Hall, Lebanon, Indiana, 2003).

3. K.H. Nierhaus, D.N. Wilson (eds). Protein Synthesis and Ribosome Structure: Translating the Genome (Wiley, Hoboken, NJ, 2004).
4. C. Branden, J. Tooze. Introduction to Protein Structure, 2nd edn (Garland Publishing, New York, NY, 1999).
5. D. Whitford. Proteins: Structure and Function (Wiley, Hoboken, NJ, 2005).
6. H.R. Nordlund, V.P. Hytönen, J. Hörnhä, J.A.E. Määttä, D.J. White, K. Halling, E.J. Porkka, J.P. Slotte, O.H. Lahtinen, M.S. Kulomaa. Biochem. J. 392 (2005) 485.
7. P.W. Rothemund. Nature 440 (2006) 297.
8. G.B. Fields, S.P. Colowick. Methods Enzymol. 289: Solid-Phase Peptide Synthesis (Academic Press, San Diego, 1997).
9. D. Bang, S.B.H. Kent. PNAS 104 (14) (2005) 5014.
10. O. Livnah, E.A. Bayer, M. Wilchek, J.L. Sussman. PNAS 90 (1993) 5076.
11. J. P. Xiong, T. Stehle, B. Diefenbach, R. Zhang, R. Dunker, D.L. Scott, A. Joachimiak, S.L. Goodman, M.A. Arnaout Sci. 294 (5541) (2001) 339.
12. L. Gilbert, J. Toivola, O. Valilehto, T. Saloniemi, C. Cunningham, D. White, A.R. Makela, E. Korhonen, M. Vuento, C. Oker-Blom. Truncated forms of viral VP2 proteins fused to EGFP assemble into fluorescent parvovirus-like particles. J. Nanobiotech. 4 (2006) 13.
13. P.A. Laurinmki, J.T. Huiskonen, D.H. Bamford, S.J. Butcher. Membrane proteins modulate the bilayer curvature in the bacterial virus Bam35. Structure. 13(12) (2005) 1819–1828.
14. V. Lucic, F. Forster, W. Baumeister. Structural studies by electron tomography: from cells to molecules. Ann. Rev. Biochem. 74 (2005) 833–865.
15. D.J. Stephens, V.J. Allan. Light microscopy techniques for live cell imaging. Science 300 (5616) (2003) 82–86.
16. D. Fotiadis, S. Scheuring, S.A. Muller, A. Engel, D.J. Muller DJ. Imaging and manipulation of biological structures with the AFM. Micron 33 (2002) 385–397.
17. M. Ren, J.A. van Kan, A.A. Bettiol, D. Lim, Y.G. Chan, B.H. Bay, H.J. Whitlow, T. Osipowicz, F. Watt. Nucl. Instrum. Methods B 260 (2007) 124.
18. H.J. Whitlow, M. Ren, J.A. van Kan, F. Watt, D. White. Nucl. Instrum. Methods B 260 (2007) 28.
19. IAEA-TECHDOC-1560. Dose Reporting in Ion Beam Therapy. (International Atomic Energy Agency, Vienna, 2007).
20. K.J. Kirkby, G.W. Grime, R.P. Webb, N.F. Kirkby, M. Folkard, K.M. Prise, B. Vojnovic, A scanning focused vertical ion nanobeam: a new UK facility for cell irradiation and analysis (2006). Nucl. Instrum. Methods Phys. Res. B 260 (2007) 97–100.
21. G. Dollinger, A. Bergmaier, A. Hauptner, S. Dietzel, G.A. Drexler, C. Greubel, V. Hable, P. Reichart, R. Krücken, T. Cremer, A.A. Friedl. Nucl. Instrum. Methods. B 249 (2006) 270–277.
22. S. Anuntalabhochai, R. Chandej, B. Phanchaisri, L.D. Yu, T. Vilaithong. Appl. Phys. Lett. 78 (2001) 2393.
23. D.J. White, S. Puranen, M.S. Johnson, J. Heino. The collagen receptor subfamily of the integrins. Int J Biochem Cell Biol. 36(8) (2004) 1405–1410.
24. I.G. Brown, K.A. Bjornstad, E.A. Blakely, J.E. Galvin, O.R. Monteiro, S. Sangyuenyongpipat. Plasma Phys. Control. Fusion 45 (2003) 547–554.
25. B. Alberts, D. Bray, K. Hopkin, A. Johnston, J. Lewis, M. Raff, K. Roberts, P. Walter. Essential Cell Biology (Garland Science, New York, 2004) p. 596.
26. S. Gorelick, N. Puttaraksa, T. Sajavaara, S. Singkarat, H.J. Whitlow, Microfluidic device fabrication using MeV ion beam Programmable Proximity Aperture Lithography (PPAL). Nucl. Instrum. Methods B 266 (2008) 2461.
27. S. Gorelick, P. Rahkila, A. Sagari A.R, T. Sajavaara, S. Cheng, L.B. Karlsson, J.A. van Kan, H.J. Whitlow. Nucl. Instrum. Methods B 260 (2007) 130.
28. M.J. Madou. *Fundamentals of Microfabrication: The Science of Miniaturisation*, 2nd edn (CRC Press, Boca Raton, 2002) pp. 562–569.

29. M.J. Madou. *Fundamentals of Microfabrication: The Science of Miniaturisation*, 2nd edn (CRC Press, Boca Raton, 2002) pp. 584–585.
30. M. Ueda, M.M. Silva, C. Otani, H. Reuter, M. Yatsuka, C.M. Lepienski, L.A. Berni. Surf. Coat. Technol. 169–170 (2003) 408.
31. P. Sioshansi, R.W. Oliver, F.D. Mathhews. J. Vac. Sci. Tech. A 3 (1985) 2670.
32. A. Dambreville, J.G. Roland. J. Bone Joint Surg. Br. 84-B (2002) Supp-I, 53.

High Speed Electronics

Mikael Östling and B. Gunnar Malm

1 Introduction

There is no question about the tremendous impact of the micro- and nanoelectronic technology on the world economy and welfare. Since the introduction of the integrated circuit about 40 years ago, the performance has followed a steady scaling outlined already in 1965 by Gordon Moore [1]. What is fascinating to note is that the most commonly produced human artefact today is the transistor. Each year we produce more than 10^{18} transistors, which is more than all printed characters each year in all books, magazines, and newspapers in the world. Most amazingly, the price per transistor is lower than the price per printed character in common magazines. We also see that this furious development will continue with unbroken momentum at least another decade. Market forecasts indicate that the gross world product in 2010 will be approximately \$1000 billion.

The semiconductor device technology has met many sceptical comments about hitting the final roadblocks. Every time, new solutions have been presented, and even more revolutionary fabrication and device design methods. Are we close to meeting a fundamental limit as the technology approaches the 22 nm node? Are we running out of atoms? Let's assume a transistor with a gate length of 22 nm and a width of 100 nm; the number of dopant atoms is around ten! Is this possible to control? Yes indeed, since the carrier transport occurs in a slab of inverted charges controlled by the gate potential [2–4].

1.1 Technology Trends for High Speed Electronics

In this section we will present technology trends for high speed electronics, with focus on silicon technology, and discuss the predictions in the International

M. Östling (✉)
School of Information and Communication Technology (ICT), KTH – Royal Institute of Technology, Kista SE-16440, Sweden
e-mail: ostling@kth.se, gunta@kth.se

Technology Roadmap for Semiconductors (for short, ITRS roadmap). The performance of both complementary metal oxide semiconductor (CMOS) and silicon-germanium (SiGe) bipolar/CMOS (BiCMOS) [5] has been significantly improved in the last 5–10 years. CMOS is now the main choice for radio frequency (RF) integrated circuits, and SiGe BiCMOS has demonstrated record performance in terms of gate delay and cut off frequencies and has recently reached into the half-terahertz domain [6, 7]. In fact, SiGe devices have similar performance as traditional high-speed devices in III–V semiconductor materials, such as InP based heterojunction bipolar transistors (HBTs) or high electron mobility transistors (HEMTs). This can be explained by improved materials—carbon doped SiGe, self-aligned technology with minimized parasitic resistances and capacitances, use of advanced lithography, etc. The evolution of SiGe BiCMOS over four different technology generations is illustrated in Fig. 1.

For CMOS, the gate length scaling is the driving force, which results in improved RF performance. Other improvements such as strained silicon channels, with high carrier mobility, also benefit the RF performance because of the higher intrinsic gain of the MOSFET. Issues related to the voltage scaling are a concern, since a supply voltage of less than 1 V makes it difficult to realize enough voltage swing in an RF/analog circuit. The reliability and breakdown of the thin gate dielectric are also limiting for the RF potential of a highly scaled CMOS. One important advantage of the III–V materials, compared to silicon-based technology, are the well developed optical components such as tunable lasers, which are monolithically integrated with photo detectors and high speed modulators [8]. However, significant progress has been made in silicon CMOS based optoelectronics recently, and an interesting option is hybrid integration of InP, by bonding or epitaxial growth, into the baseline CMOS process. In Sect. 3 below we will discuss some performance boosters such as different types of substrates SOI, and virtual substrates. Strain enhanced mobility is
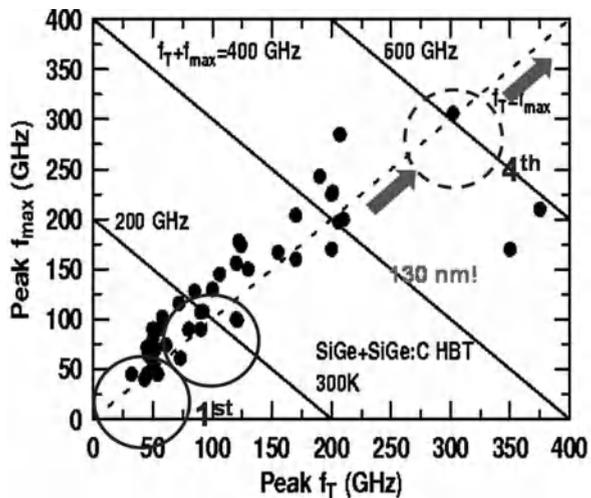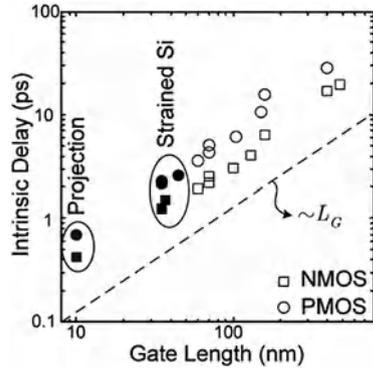

Fig. 1 Example of SiGe performance evolution over four technology generations. Courtesy of John D. Cressler, Institute of Technology, Atlanta, GA [5]

Fig. 2 CMOS performance
trends vs. gate length scaling,
showing the introduction of
strained silicon. Reproduced
after Antoniadis et al. [10]
with permission from IBM
Journal of Research and
Development

a very important concept to keep CMOS on the roadmap [9]. Figure 2 illustrates how
strain induced mobility enhancement was introduced at about the 90 nm technology
node (see the next section for a definition), to ensure continued performance scaling
for shorter gate lengths.

To reduce the increasing gate leakage current a paradigm shift, in which the sili-
con dioxide gate insulator is replaced by a high-κ/metal gate stack was necessary. It
should be mentioned that the use of ion beam methods for characterising the κ/metal
gate stack is discussed further in chapters "Low and Medium Energy Ion Scattering
for Near Surface Structure and Nanoscale Depth Profiling" and "Thin Film Charac-
terisation by Means of MeV Ion Beams". For coming technology nodes, advances
in junction technology, both with regard to dopant introduction and to activation,
will be crucial.

1.2 International Technology Roadmap for Semiconductors

To illustrate some current trends, we will start by analyzing the predictions by the
ITRS roadmap [11]. Here technology evolution is predicted and current roadblocks
are highlighted to direct research efforts. The roadmap outlines the key requirements
for coming technology generations. In particular, ion implantation (I/I) is covered
in the section on front-end process technology and described with device-related
parameters such as junction depth (X_j) and extension sheet resistance (R_S). Some
difficult challenges are outlined in Table 1 below:

The roadmap uses a metric which refers to the minimum half-pitch of a specific
feature, e.g., contacted gates. This is referred to as the technology node. Hence, the
roadmap does not define a target gate length as such. Currently, 45 nm technology
has been introduced, and the roadmap predictions focus on the upcoming 32 and
22 nm nodes, which are in development. A major paradigm shift took place as high-
κ/metal gate was successfully introduced at the 45 nm node. The roadmap states
that further reduction of the high-κ dielectric thickness is one of the objectives for
coming technology nodes. For the physical gate length, improved control of the

Table 1 Example of front-end process difficult challenges from ITRS 07

Difficult challenges ≥ 22 nm	Summary of issues
Thermal/thin films/doping/etch	Introduction of high-κ/metal gate into high performance and low operating/low standby power and equivalent oxide thickness (EOT) scaling below 0.8 nm
	Scaling extension junction depths below 10 nm while achieving high dopant activation
	Gate critical dimension control for physical gate length < 20 nm

critical dimension is required, which includes improved tolerances in the combined lithography and etching steps as well as metrology techniques to measure 20 nm features across a large diameter wafer.

Finally, severe challenges for I/I are anticipated. Achieving 10 nm X_j along with high dopant activation requires very low energy (less than 1 keV) implants and a thermal process which maximizes the dopant activitation, with little or no diffusion. As seen from Table 2 the scaling of X_j will continue, and on a five-year horizon no known solutions are in sight. The same applies to the sheet resistance of the activated extension dopants.

The junction depth issue is controlled by lowering the implantation energy, while the activation is controlled by pre-amorphization or co-implantation of different dopants species. The complexity of the junction formation is therefore significantly increased; a single implanation of one species such as As or B is not sufficient. Ge is used for pre-amorphization and C or F are suggested for co-implantations. There is much interest in diffusionless activation techniques. Among the candidates are laser annealing and solid phase epitaxial regrowth (SPE). In general, annealing times have been reduced, from rapid thermal anneal (RTA)—duration typically 10 s, to spike anneal a fast jump to high temperature (1050–1100 °C, duration 1 s)—immediately followed by fast cool down, and finally millisecond anneal (MSA).

Table 2 Example of predictions in the ITRS roadmap related to dopant introduction

• Year of production	2008	2009	2010	2011	2012
Drain extension Xj (nm)	7.5	7	6.5	5.8	4.5
Maximum allowable parasitic series resistance × width (Ω−μm)	140	120	105	80	70
Maximum drain extension sheet resistance (Ω/sq)	740	677	650	548	593
Extension lateral abruptness for bulk (nm/decade)	2.5	2.2	2	1.8	1.5
Contact Xj (nm)	25.3	22	19.8	17.6	15.4

Color coding: *white/light* and *gray* = solution known, *dark gray* = no known solution. The values refer to bulk MPU/ASIC technology and were obtained from the ITRS 06 update [11]

2 Ion Implantation in High-Speed Device and Integrated Circuit Process Technology

This section outlines the use of I/I in high-speed device and very large scale integrated (VLSI) circuit process technology with a focus on silicon CMOS devices and circuits.

Since the early 1980s, ion implantation has been one of the reliable workhorse processes in the VLSI technology. A schematic view in Fig. 3 below shows the various areas in a CMOS device process in which I/I plays an important role, mainly for the channel and contact formation, but also for the background doping level in the so-called wells.

The number of I/I steps in a state-of-the-art process is more than 20. The various different fluencies and energies needed are represented in Fig. 4.

Note that fluence is the SI unit corresponding to the number of ions impinging per unit area. In the semiconductor industry the standard terminology is "dose", which in other applications has the meaning of deposited energy per kg. The dose range is from 10^{11} to 10^{16} cm^{-2} for device fabrication, whereas higher doses are needed for some special applications such as SOI substrates; see Sect. 3.3 below. The I/I energy range is from several hundreds of eV for ultra shallow junctions to several MeV for retrograded wells; see Fig. 3. It can be inferred from this discussion that several dedicated tools are needed to cover this energy-dose space. The tools are normally grouped into low- and high-energy tools, respectively. The implantation current
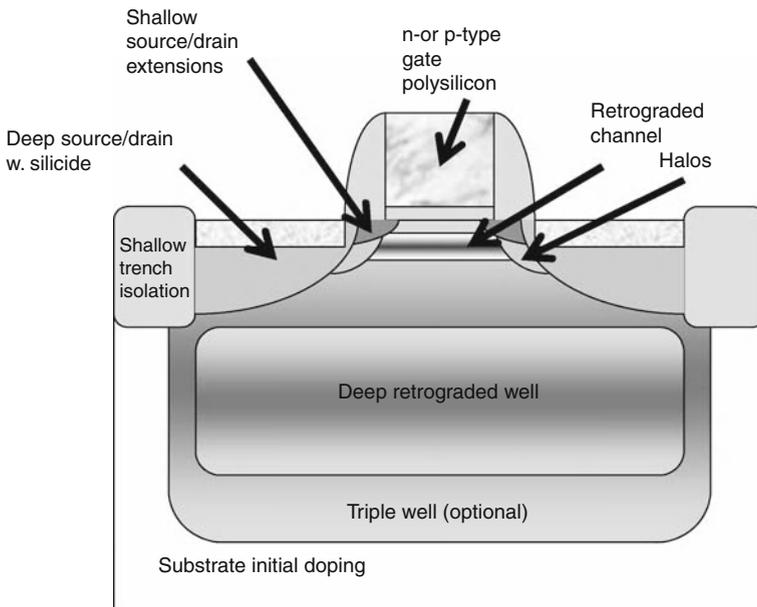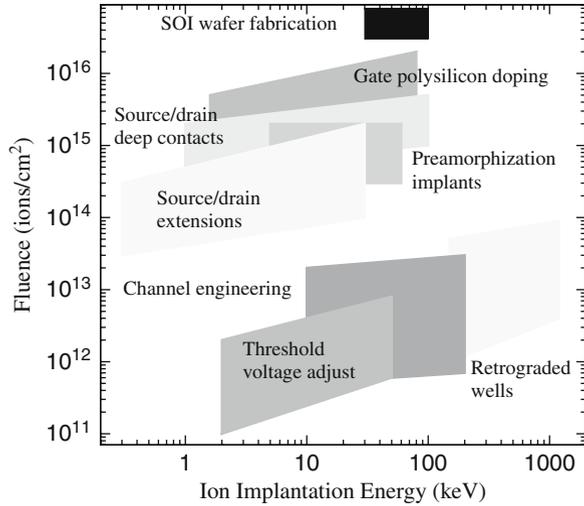
Fig. 3 Overview of the I/I steps needed to fabricate integrated circuits in CMOS technology

Fig. 4 Fluence-energy space
of ion implantations for
silicon device and process
technology. Adapted after
Doering and Nishi [12]

is the second important category; however in view of maximum wafer through-put, there is limited use for low current. Special configurations are used if angled implants (tilting) or wafer rotation is needed. Single wafer tools are predominant in the industry because of uniformity constraints.

In this context, it should be mentioned that CMOS uses two type of transistors, n-channel and p-channel field effect transistors (FETs). The basic building block in a circuit is the inverter, consisting of one pFET and one nFET. To form pFETs and nFETs in the same substrate, separate regions of opposite doping polarity are needed. For the n-type doping As is preferred because of its slower diffusion compared to P; an exception to this rule is of course deep diffusions or buried layers. For the p-type doping B is the only candidate with high enough solubility and hence enough active dopants. To control the depth of the profile, molecular ions or clusters such as BF_2 and $B_{10}H_{14}$ can be used. However, BF_2 has a disadvantage, in that the presence of fluorine counteracts high activation of B. The p-type and n-type substrate regions are achieved by high energy I/I of wells, sometimes called tubs. By using high energy implantation the peak of the well profile can be positioned without long thermal diffusion of dopants; this gives a good control over lateral dimensions, and the device-device separation and footprint can be reduced. A so-called retrograde profile, in which the peak is located at the bottom of the well, is desired to suppress so-called latch up, which is an undesired leakage path between n-well and p-well. The commonly used structures today are referred to as twin well or triple well, of which the latter provides independent body biasing and suppressed substrate noise in mixed signal applications. Well formation is normally self-aligned to the already formed oxide-filled shallow trench isolation (STI).

Doping of the channel region requires several implantations of varying energy and dose. Each has its specific purpose, namely punch through stopper, retrograde channel implant, and threshold voltage (V_T) adjust. See Fig. 5 for typical depth

Fig. 5 Typical doping depth
profile in the channel and
well regions of a 50 nm gate
length MOSFET, based on
simulation data
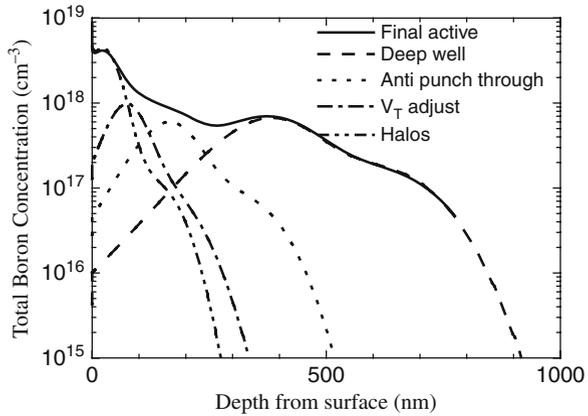
profiles, including the retrograde p-well. The dose variation control of the V_T adjust
is critical for device and circuit applications.

The source and drain regions are formed in a self-aligned manner, after the pat-
terning of the gate electrode. The source/drain implantations include a deep contact
region with low sheet- and contact-resistance and a shallow region, which is called
the extension. Sidewall spacers are used to offset the deep region from the intrinsic
transistor channel.

At this step in the process sequence, the dual doping of the polysilicon gates is
also introduced by masked ion implantations.

High activation and shallow junctions are needed to optimize short channel
effects and to minimize the junction capacitance. These are two competing require-
ments, and they are intimately coupled by the point defect creation and diffusion.
Implant energies of 1 keV or less are used for the B-profiles in the pFETs.

As the implant energies are reduced further, it becomes difficult to create an ion
beam with high enough current density and energy purity. A future potential tech-
nique for the purpose of low energy, high current implantation is the plasma immer-
sion ion implantation technology. This technique was researched heavily during the
early 1990s and successful demonstrations have been published recently [13–15].
Commercial applications have been found in highly doped polysilicon for memory
applications.

Halo implants, as shown in Fig. 6, can be used to counterdope the extensions and
hence make them even more abrupt [16]. The halo implantations are normally per-
formed with a tilt of 30–60°. Using a tilt puts some of the halo dose in the channel,
and hence V_T is affected. Especially for short channel devices, the halos or pockets
are present everywhere in the channels, i.e., the profiles from the source and drain
side, respectively, will merge. This gives rise to a phenomenon known as reverse
short channel effect. Tilted implantations of 7° are commonly used to avoid boron
channelling in preferred crystal directions.

Alternative techniques for the activation anneal such as spike or laser annealing
should be considered. Spike annealing takes advantage of the different time con-

Fig. 6 Example of tilted ion implantation to form the so-called HALO regions to prevent short channel degradation in ultra-scaled MOSFETs
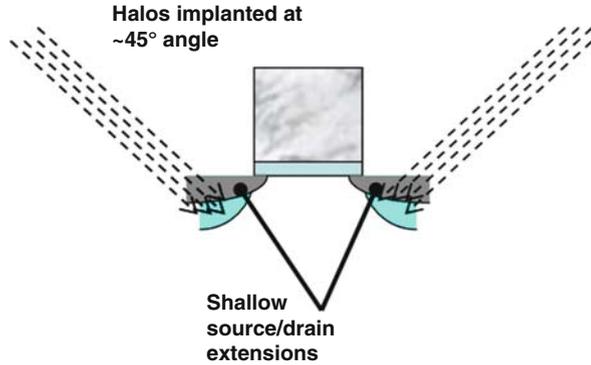
stants involved in defect and dopant diffusion, and laser annealing causes a local recrystallization of the doped region. Activation of shallow junctions can be further improved by *pre-amorphization* using Ge ion implantation or *co-implantation* using C or F [17–20] with different dopant species, especially B. Both these methods affect the end of range implantation defect profiles and hence give additional control of the final resulting active profile.

3 Performance Boosters in CMOS Technology

The need for ever higher device performance is very challenging. In fact, the need for better materials properties has led to an increased effort in materials technology. In this section both substrate material engineering and new material choices will be discussed.

3.1 High-κ/Metal Gate

From the 45 nm node the standard SiO_2 gate dielectric cannot fulfil the requirements in the roadmap. This is especially important for low-power operation, and high-κ/metal gate has not yet been optimized for high-speed devices. The first announcement from the industry of products incorporating high-κ/metal gate was made in 2007, shortly before the ITRS roadmap prediction. These devices use hafnium Hf-based stacks, including well controlled interfaces to both silicon channel and gate electrode. In fact, the industry standard polysilicon gate technology is not compatible with Hf-based films unless intermediate layers are used.

From a fundamental device physics view, several candidate materials have been studied and Hf offers a medium κ-value in combination with suitable conduction and valence band offsets for low direct leakage. For ultimate technology scaling, other rare earth metals must be considered, such as lanthanum, La [21–23].

Ion beam techniques have found wide use in the structural analysis of high-k materials. In chapter "Low and Medium Energy Ion Scattering for Near Surface

Fig. 7 RBS analysis of 30 nm high-k film to determine (stoichiometric) composition. Hf : O = 1 : 2.08, Zr: 1.5%, Ar: 1%
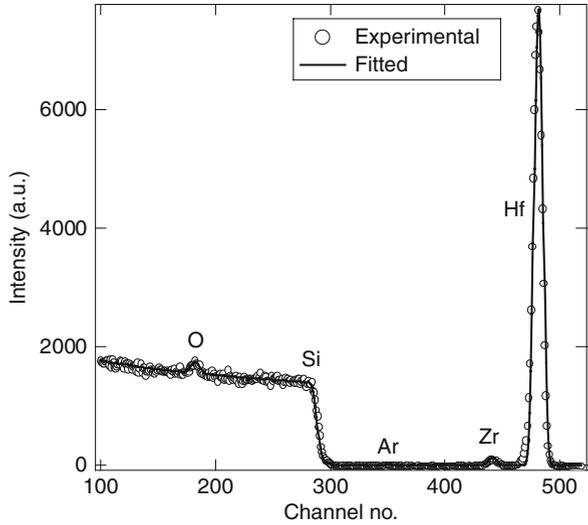
Structure and Nanoscale Depth Profiling" *medium energy ion scattering* and *time of flight-energy elastic recoil detection analysis* (ToF-E ERDA) will be discussed. Another commonly used technique is *Rutherford Back Scattering* (RBS). RBS is an important analysis tool for thin films in microelectronics. The compositon of Hf-based films can be determined with resolution below the percent level; an example is shown in Fig. 7 [24].

3.2 Substrate Engineering

Improved performance of standard CMOS technology is possible by replacing the bulk silicon wafer with another starting material. Two types of substrates will be discussed, silicon-on-insulator (SOI) and strained virtual substrates (VS). The two main advantages of SOI are reduced junction-substrate leakage, and hence low power operation; and better control of the short-channel effects, since the buried oxide layer prevents the drain side electric field from reaching the source. Using VS technology the fundamental properties of silicon, mainly carrier mobility, can be improved. A more unconventional solution would be to use high mobility III-V material in the channel region. This is currently a subject of ongoing research and will be mentioned below.

3.3 SOI

SOI was originally intended for low-power logic and also used for a long time in radiation-hard circuits. However, at the $0.18\,\mu$m CMOS technology node,

SOI was introduced in microprocessors by some companies to increase the performance/speed.

Ultrathin body SOI is one of the technologies on the roadmap beyond the current 45 nm node. However, the introduction of ultrathin body SOI has been pushed back to later nodes, because of innovations in other areas.

In this section we will discuss the fabrication of high quality SOI substrates using the most common techniques, namely separation by implantation of oxygen (SIMOX) and Smartcut®/Unibond®. Ion implantation is one of the key technologies to realize this type of substrate. Also needed are wafer bonding and planarization. In SIMOX wafers, a buried oxide layer is formed by a high dose (fluence) $(1 \times 10^{18}\,\mathrm{cm}^{-2})$ followed by a high temperature $(1320°C)$ anneal. For reasonable throughput, high current implanters must be used with a beam current of $100\,\mathrm{mA}$. This is beyond the capability of standard implanters. On the other hand, SmartCut® relies on hydrogen implantation of an oxidized wafer to a dose (fluence) of about $5 \times 10^{16}\,\mathrm{cm}^{-2}$; cf. Fig. 4 above. A schematic process flow is illustrated in Fig. 8. The implanted hydrogen creates local voids or microcavities in the wafer used as starting material. This wafer is then bonded to another oxidized handle wafer and annealed at $600°C$ to produce a cut (at the peak of the hydrogen profile). The result becomes one SOI wafer with a thin, device-grade, silicon top layer, and one handle wafer for reuse.
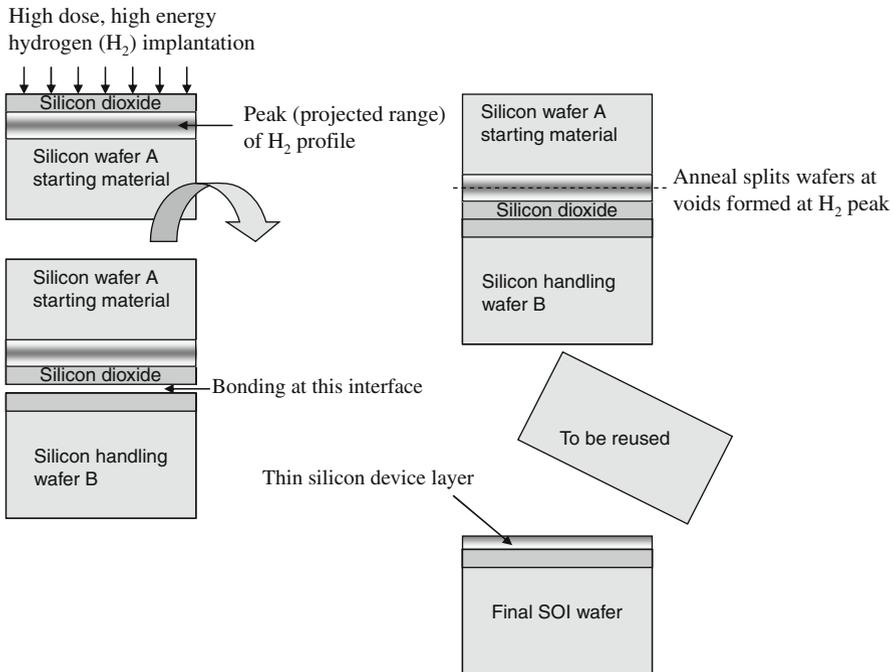

Fig. 8 Example of hydrogen implantation to form an SOI substrate with a buried oxide layer [16, 25]

3.4 Virtual Substrates

Although silicon has many desirable properties for integrated circuit manufacturing, the intrinsic carrier mobility for both electrons and holes is comparatively low. A strained silicon layer can be utilized to significantly increase the mobility in the MOSFET channel. One way to accomplish this is the use of biaxially strained virtual substrates (VS). The other approach is locally induced uniaxial strain, using dielectric films as stressors. In VS technology, strain is introduced by gradually changing the lattice constant by introducing high concentrations of Ge. Typically 1% of Ge is added per micrometer, until a final concentration of 20–30% is reached, e.g., a composition of $Si_{0.8}Ge_{0.2}$. This constitutes the SiGe relaxed buffer layer. On top of this layer a thin silicon layer is grown. The lattice mismatch between the silicon and the relaxed $Si_{1-x}Ge_x$ results in a tensile strain in the thin silicon films, since Ge has a larger lattice constant. A high resolution TEM image of a VS substrate is shown in Fig. 9.

3.5 Heteroepitaxy

What will be the technology roadmap after 2015? The European technology platform ENIAC has depicted the scenario in their recently published Strategic Research Agenda [26]. The main scenario in the agenda is still based on a silicon platform; however, some add-on solutions are pointing to the introduction of III–V technology on silicon.

Concerning enhanced speed, many options are already approaching enough maturity to be applied in commercial technology. Examples include inducing stress in the conductive channel of the transistor and/or using different crystal orientations. In a longer term scenario, increased speed may be achieved by replacing silicon channels with more conductive materials such as germanium (Ge) or III-V

Fig. 9 High resolution transmission electron microscope (TEM) image (with near atomic resolution) of a complete MOSFET device structure on a virtual substrate. The labels indicate the gate oxide (SiO_2), polysilicon gate electrode (p-Si), strained channel layer (Si), and relaxed SiGe buffer layer
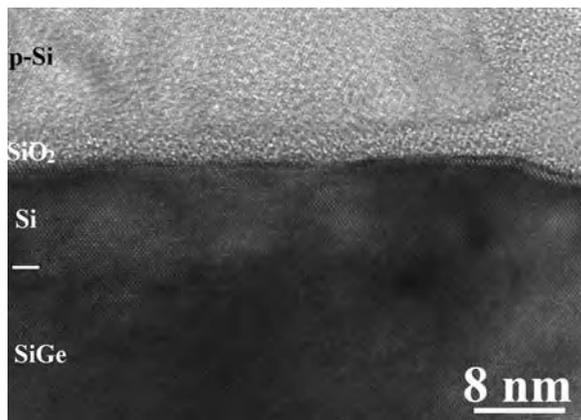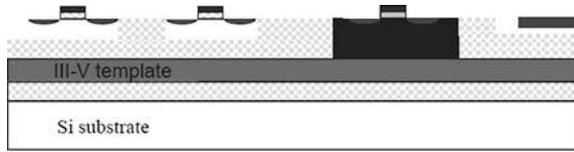
Fig. 10 Cross section of a future silicon wafer with integrated areas of III-V devices. Courtesy of E. Fitzgerald, Massachusetts Institute of Technology, Cambridge

compounds embedded in more sophisticated architectures (e.g., germanium-on-insulator I, III–V on germanium, and nanowires) [26].

By using lattice matching, other high mobility materials such as GaAs [27, 28] and InSb [29] compounds can be used in the channel of a field effect transistor. Excellent results have been achieved with zero antiphase boundary GaAs grown on Si. Introducing a clever materials platform called SOLES (silicon on lattice engineered substrates) MIT researchers have shown that this integration technique also is manufacturable. The concept is schematically illustrated in Fig. 10, in which a buried III-V template layer allows growth in selected device regions, while silicon is available on the remaining part of the wafer surface to form more conventional devices.

4 Lithography

In this section we review current trends in lithography for CMOS integrated circuit applications. The patterning of the small features in integrated circuits is one of the increasingly difficult challenges as the minimum pitch as well as the isolated feature size continue to shrink with scaling. The feature size is now considerably smaller than the optical wavelength used in lithography. In fact, standard DUV lithography with a 193 nm light source is used for the 45 nm CMOS technology node in which the patterned gate length is even smaller—about 30 nm.

This has been made possible because of different kinds of improvements of the lithographic process, such as phase shifting masks (PSM) and optical proximity correction (OPC). However, the underlying principle is a detailed modelling of the electromagnetic wave propagation in the optical projection system, including lenses and the reticle itself. In this way the reticle can be corrected to produce the desired shape of the printed feature on the wafer. A common measure of the lithographic capability is the area of the very dense 6T-SRAM memory cell, which today occupies less than $0.3\,\mu m^2$.

As the use of the 193 nm tools reaches its limitation, one important technology change is to introduce a medium with higher refractive index in the optical path. Immersion in water can increase the numerical aperture (NA) of the lenses, and smaller patterns can be resolved. In practice, this means that a thin layer of water is placed between the optical projection column and the resist coated wafer surface.

Much effort is also placed on light sources with shorter wavelength, so-called extreme ultraviolet (EUV). This will be accompanied by a change in the optics which will be reflective (using mirrors) instead of refractive (lenses). Masks which are transparent to these wavelengths are also needed.

The alternatives to optical lithography using reticles are based on direct-writing techniques, in which an electron beam is widely used for nanostructures and for special applications, such as dense optical gratings. Focused ion beam (FIB) systems have also been proposed. For an extensive treatment of FIB, refer to chapter "Focused Ion Beam Machining and Deposition". Typically, ions such as Ga^+ (atomic mass 31) are used. For FIB lithography, lighter ions such as protons (H^+) may be used; for more details see chapter "Proton Beam Writing: A New 3D Nano Lithographic Technique". Some of the issues here are wafer throughput and line edge roughness due to electron/ion scattering. Electron beam systems are widely used for lithographic reticle (mask) manufacturing. FIB is used for mask repair and also for postprocessing analysis and correction of failures in integrated circuits.

Finally, we briefly discuss a technique called sidewall transfer (spacer) lithography, in which a combination of optical lithography and selective directional etching is used [30]. This technique has superior control of line edge and line width roughness compared to, e.g., electron beams. The basic principle is to create a thin sidewall inside an etched opening in a patterned support layer. The patterning of the support layer is done by optical lithography. The support layer in then removed, by selective etching, and the sidewall will now serve as a hard mask for the actual 'device pattern,' e.g., a polysilicon gate line, as shown in Fig. 11.

Fig. 11 Example of a 50 nm polysilicon gate patterned by sidewall transfer lithography, showing the effect of line width and line edge roughness (LWR/LER)
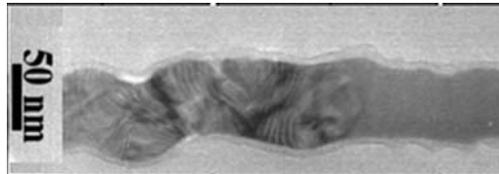
5 Conclusions

This chapter has discussed the background, present technology concerns, and short future outlook for nanoelectronic semiconductor technology. We have addressed the immense importance of the technology from a socioeconomic perspective and shown that rapid progress continues with unbroken strength. We have also discussed selected technology solutions and commented on the importance of ion implantation technologies in this field. Finally, a brief outlook on new substrate and materials requirements was given, and an alternative based on heterogeneous materials integration was shown.

References

1. Moore GE, Cramming more components onto integrated circuits, *Electronics*, **38**, 114–117, (1965).
2. Takeuchi K, Tatsumi T, and Furukawa A, Channel engineering for the reduction of random-dopant-placement-induced threshold voltage fluctuation, *Technical Digest of the International Electron Device Meeting*, 841–844, (1997).
3. Wettstein A, Penzin O, Lyumkis E, et al., Random dopant fluctuation modelling with the impedance field method, *Proceedings of International Conference on Simulation of Semiconductor Devices and Processes*, 91–94, (2003).
4. Ohtou T, Sugii N, and Hiramoto T, Impact of parameter variations and random dopant fluctuations on short-channel fully depleted SOI MOSFETS with extremely thin box, *IEEE Electron Device Letters*, **28**, 740–742, (2007).
5. Cressler JD and Niu G, *Silicon-Germanium Heterojunction Bipolar Transistors*. Norwood, MA: Artech House, 2003.
6. Chevalier P, Zerounian N, Barbalat B, et al., On the use of cryogenic measurements to investigate the potential of Si/SiGe:C HBTs for terahertz operation, *Proceedings of Bipolar/BiCMOS Circuits and Technology Meeting*, 26–29, (2007).
7. Krithivasan R, Yuan L, Cressler JD, et al., Half-terahertz operation of SiGe HBTs, *IEEE Electron Device Letters*, **27**, 567–569, (2006).
8. Westergren U, Yu Y, and Thylen L, High-speed travelling-wave electro-absorption modulators, *Proceedings of SPIE – The International Society for Optical Engineering*, **6350**, 635004, (2006).
9. Thompson SE, Armstrong M, Auth C, et al., A logic nanotechnology featuring strained-silicon, *IEEE Electron Device Letters*, **25**, 191–193, (2004).
10. Antoniadis DA, Aberg I, NíChléirigh C, et al., Continuous MOSFET performance increase with device scaling: The role of strain and channel material innovations, *IBM Journal of Research and Development*, **50**, 363–376A, (2006).
11. International Technology Roadmap for Semiconductors (ITRS) 2006 update, 2006.
12. Doering R and Nishi Y, *Handbook of Semiconductor Manufacturing Technology*, 2nd ed. Boca Raton, FL: CRC Press, Taylor & Francis Group, 2008.
13. Kilho L, Plasma immersion ion implantation as an alternative doping technology for ULSI, *Extended Abstracts of the Second International Workshop on Junction Technology*, 21–27, (2001).
14. Agarwal A and Kushner MJ, Characteristics of pulsed plasma doping (P^2LAD) sources for ultra-shallow junction formation, *Proceedings of IEEE International Conference on Plasma Science*, 14, (2005).
15. Current MI, Plasma immersion implantation for modification and doping of semiconductor materials: A historical perspective: 1886 to 2000, *Extended Abstracts of the First International Workshop on Junction Technology*, 13–17, (2000).
16. Chen W-K, *The VLSI Handbook*, 2nd ed. Boca Raton, FL: CRC Press, Taylor & Francis Group, 2006.
17. Augendre E, Pawlak BJ, Kubicek S, et al., Superior N- and PMOSFET scalability using carbon co-implantation and spike annealing, *Proceedings of 36th European Solid-State Device Research Conference*, 355–358, (2006).
18. Yamamoto T, Kubo T, Sukegawa T, et al., Advanced junction profile engineering featuring laser spike annealing and co-implantation for sub-30-nm strained CMOS devices, *Technical Digest of Symposium on VLSI Technology* 2, (2006).
19. Kubicek S, Hoffmann T, Augendre E, et al., Superior nMOSFET scalability using Fluorineine co-implantation and spike annealing, *Proceedings of International Conference on Advanced Semiconductor Devices and Microsystems*, 101–104, (2006).

20. Pawlak B, Lindsay R, Surdeanu R, et al., Optimizing p-type ultra-shallow junctions for the 65 nm CMOS technology node, *14th International Conference on Ion Implantation Technology Proceedings*, 21–24, (2003).
21. Robertson J, Band structures and band offsets of high K dielectrics on Si, *Applied Surface Science*, **190**, 2–10, (2002).
22. Iwai H, Ohmi S, Akama S, et al., Advanced gate dielectric materials for sub-100 nm CMOS, *Technical Digest of International Electron Device Meeting*, 625–628, (2002).
23. Engström O, Raeissi B, Hall S, et al., Navigation aids in the search for future high-*k* dielectrics: Physical and electrical trends, *Solid-State Electronics*, **51**, 622–626, (2007).
24. Malm BG, Sjöblom G, Martin D, et al., *Unpublished Results*, (2007).
25. Wolf S and Tauber RN, *Silicon Processing for the VLSI era Volume 1 – Process Technology*, 2nd ed. Sunset Beach, CA: Lattice Press, 2000.
26. Eniac strategic research agenda, http://www.eniac.eu/web/downloads/SRA2007.pdf, 2007.
27. Fitzgerald EA, Dohrman CL, Chilukuri K, et al., Epitaxial growth of heterovalent GaAs/Ge and applications in III–V monolithic integration on Si substrates, *ECS Transactions, SiGe and Ge: Materials, Processing, and Devices*, **3**, 561–568, (2006).
28. Carlin JA, Andre CL, Kwon O, et al., III–V device integration on silicon via metamorphic SiGe substrates, *ECS Transactions, SiGe and Ge: Materials, Processing, and Devices*, **3**, 729–743, (2006).
29. Datta S, Ashley T, Brask J, et al., 85 nm gate length enhancement and depletion mode InSb quantum well transistors for ultra high speed and very low power digital logic applications, *Technical Digest of International Electron Devices Meeting*, 763–766, (2005).
30. Hållstedt J, Hellström P-E, Zhang Z, et al., A robust spacer gate process for deca-nanometer high-frequency MOSFETs, *Microelectronic Engineering*, **83**, 434–439, (2006).

Surface Modification Using Reactive Landing of Mass-Selected Ions

**Peng Wang and Julia Laskin**

1 Introduction

Collisions of ions with surfaces play an important role in a variety of scientific disciplines, including surface science, materials science, mass spectrometry, imaging, and spectroscopy. Ion-surface collision phenomena are broadly used in many analytical methods and surface modification techniques, including secondary ion mass spectrometry (SIMS) [1–3], chemical imaging [4], ion scattering spectrometry [5], surface-induced dissociation (SID) [6–8], ion soft landing [9, 10], growth and modification of thin films [11], and formation of three-dimensional nanostructures [12]. Physical phenomena that occur during ion-surface collisions have been extensively reviewed [9–11]. These include physisorption or ion deposition, elastic and inelastic scattering, chemical reactions, charge exchange, sputtering, and ion implantation. The competition between these processes is determined by the kinetic energy of the projectile ion and the properties of the surface. Different regimes of kinetic energies schematically shown in Fig. 1 are approximately defined as follows: thermal (below 1 eV), hyperthermal (1–100 eV), low-energy (0.1–10 keV), medium-energy (10–500 keV), and high-energy (>500 keV) [5].

This review will focus on phenomena that occur during collisions of polyatomic ions with surfaces at hyperthermal energies, with specific emphasis on bond making and bond breaking processes. Because in this energy regime the initial kinetic energy of the ion is commonly sufficient for breaking chemical bonds on the surface but is too low for significant penetration of a polyatomic ion into the surface, hyperthermal collisions are well suited for chemical modification of the outer layer of the surface without substantial perturbation of the bulk substrate. In addition, careful control of the properties of the ion beam can be used for very specific surface modification. Fundamental investigation of collision phenomena in this energy

J. Laskin (✉)

Fundamental Science Division, Pacific Northwest National Laboratory, Richland, WA 99352, USA
e-mail: Julia.Laskin@pnl.gov

Fig. 1 Fundamental physical phenomena during ion-surface collision processes occurring at different energy regimes
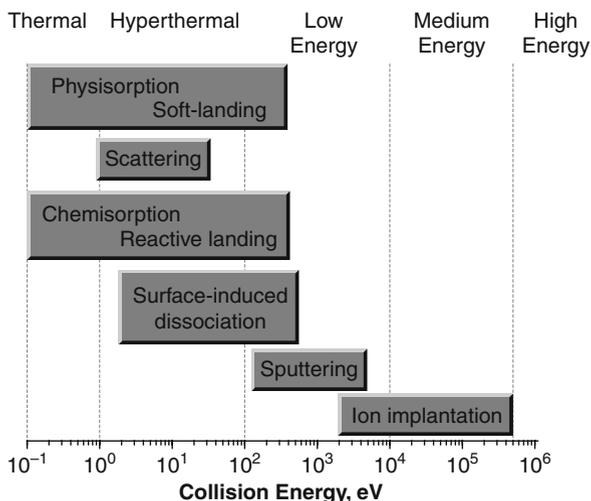
regime is important for the development of new approaches in plasma processing of substrates, understanding plasma-wall interactions and surface corrosion in fusion reactors, development of new sensors and biosensors, and development of new biomaterials for a variety of biological applications. Mass- and kinetic-energy selected ion beams are ideal model systems for studying complex processes related to physical and chemical modification of surfaces by hyperthermal ions in plasma processing and other applications.

The chapter is organized as follows. We start by discussing physical and chemical processes associated with collisions of hyperthermal ions with surfaces, and briefly describe instrumentation used in mass-selected ion deposition experiments and surface characterization approaches. We then present several examples of surface modification using hyperthermal ions, describe factors that affect the efficiency of ion-surface reactions, and briefly discuss future directions in this area of research.

2 Collision Processes at Hyperthermal Energies

Physical and chemical processes occurring during interaction of hyperthermal polyatomic ions with surfaces are summarized in Fig. 2 [9]. These include (i) simple elastic scattering of ions, (ii) surface-induced dissociation, (iii) reactive ion-induced desorption or chemical sputtering, (iv) reactive collisions resulting in chemical transformation of the ion, the surface, or both, (v) charge reduction, neutralization, and charge inversion, and (vi) ion deposition or soft landing. For detailed discussion of these phenomena, the reader is referred to a number of excellent reviews [7, 10, 11, 13].

Scattering and deposition of ions are the major competing processes in the hyperthermal energy regime. Deposition of ions on surfaces will be discussed in detail in the next sections. Scattering is often accompanied by the vibrational excitation and subsequent surface-induced dissociation (SID) of the projectile ion. Fast vibrational excitation of the ion during collision with a surface is advantageous for structural characterization of molecules using tandem mass spectrometry [14–18]. In the last decade, SID became a very powerful tool for studying structures and fragmentation mechanisms of large gaseous ions 2[6, 8].

Reactive collisions of ions with surfaces result in reactive scattering, chemical sputtering, and attachment of intact projectile ions or their fragments to the surface. Reactive scattering, defined as scattering accompanied by transfer of atoms or functional groups from the surface to the projectile ion, is commonly observed for collisions of odd-electron ions with surfaces. This process provides a basis for surface characterization using reactive ion scattering spectrometry [19–21]. Reactive ion-induced desorption [5], often called "chemical sputtering" [22], is used to remove molecules from the outer layer on the surface by reacting them with

projectile ions. The resulting products are desorbed from the surface if their binding energy to the substrate is sufficiently low. Chemical sputtering—the basic process in plasma etching—is usually observed at much lower collision energies than physical sputtering that involves collision cascade in the solid generated by the energetic collision. It is a promising method for studying interfacial phenomena in the condensed phase [23].

Finally, we should note that charge transfer commonly competes with ion-surface collision phenomena discussed earlier. Neutralization of projectile ions by electron transfer to or from the surface efficiently competes with scattering. The extent of neutralization is substantially reduced when conductive substrates are coated with thin films, e.g., self-assembled monolayers (SAMs), Langmuir-Blodgett films, or inorganic insulating films. Charge inversion, in which the projectile ion changes its polarity, has been reported for scattering of some projectile ions off solid targets [15, 22–26]. This process is particularly important for collisions of small keV ions with substrates. However, charge inversion has also been observed for several polyatomic ions at hyperthermal energies [25]. Charge inversion cross section is strongly dependent on the nature of the surface and the collision energy of the projectile ions. Charge transfer phenomena are also important in chemical sputtering that is also commonly followed by re-ionization of desorbed species.

2.1 Soft Landing

In this review we will discuss soft and reactive landing of mass selected ions on surfaces. These processes are hyperthermal analogs of physisorption and chemisorption, respectively. Soft landing (SL), defined as deposition of intact ions onto solid or liquid surfaces, is a process that efficiently competes with ion scattering in the hyperthermal energy regime. The term "soft landing" has been used to describe two distinct processes, one in which neutralization occurs during ion-surface collision and one in which the ion preserves its charge. The soft landing phenomenon was first described in 1977 by Cooks and coworkers for collisions of small sulfur-containing ions with metal surfaces [27], while charge retention was demonstrated two decades later for collisions of small polyatomic ions with fluorinated self-assembled monolayer surfaces (FSAM) [28]. It is generally recognized that the soft landing process is favored at low collision energies. At higher energy, crash-landing, in which the ion undergoes fast fragmentation during the collision and the resulting fragments are retained on the surface, may compete with ion deposition. Soft landing found its applications in a variety of research areas, including purification of compounds from complex mixtures [29–31], deposition of mass-selected cluster ions on substrates [32–36, 38], preparation of protein or peptide arrays [39–43], preparation of novel synthetic materials [44–47], and investigation of transport properties of small ions through thin and thick films [48, 49]. More recently, soft landing has been utilized for controlled preparation of new types of nanoscale materials, e.g., nanocatalysts [50], nanomagnetics [35, 51, 52], nanoelectronics [45, 47], and biological

microarrays [42]. It has been demonstrated that ionized large biomolecules, such as DNA [53], proteins [41–43], and even whole viruses [54], can be nondestructively landed onto solid or liquid matrices and recovered with retention of biological activity. The ability to select the mass-to-charge ratio of the precursor ion, its kinetic energy and charge state, along with precise control of the size, shape, and position of the ion beam on the target makes soft landing an attractive approach for surface modification.

2.2 Reactive Landing

Reactive landing, in which projectile ions or their fragments are covalently attached to the surface following deposition, is another phenomenon of interest for practical applications [55–63]. Reactive landing is a complex process involving hyperthermal ion/surface collisions and interfacial chemistry. Although the mechanism of reactive landing is not very well understood, it is a promising approach for selective surface modification. Reactive landing offers unique advantages over conventional solution phase modification methods. These include very selective preparation of the reactant using mass spectrometry that greatly simplifies the purification stage and eliminates the effect of solvent and sample contamination on the quality of the film; control of the reaction kinetics by the kinetic energy of the projectile ion that can be readily varied from 0 to several keV; and the ability to pattern the reactant on the surface using sophisticated ion optics. A number of examples of surface modification resulting from reactive landing have been presented in the literature. For instance, OH^+ and NH^+ ions have been grafted into polystyrene films [64], as have fluorine-containing projectiles [65]. Silicon nitride formation has been achieved by reactions of N^+ and N_2^+ ions with Si (100) surface [66], and diamond-like carbon films have been deposited on clean surfaces as a result of collisions of low-energy C^+ ion beams [67]. In addition, many studies have reported ion-surface modifications with polyatomic ions in which the entire projectile ion remains intact upon covalently binding with the surface. This type of surface modification will be discussed in detail in this chapter.

Physical and chemical properties of the target and the projectile ion determine the outcome of soft and reactive landing experiments. Reactivity between the ion and the surface is necessary for reactive landing, while soft landing experiments are performed using chemically inert surfaces. Self-assembled monolayer surfaces (SAMs) are excellent deposition targets for soft landing because of their ability to dissipate the kinetic energy of the projectile and their efficiency in trapping captured species [28, 40, 51]. More importantly, SAMs also provide a simple and convenient platform for tailoring chemical properties of surfaces [68]. For example, CH_3- and CF_3-terminated SAMs are fairly inert surfaces perfectly suited for studying soft landing phenomena, while SAM surfaces terminated with amine, hydroxyl, carboxylic acid, phosphate, aldehyde, and halogen groups are susceptible to a variety of reactions, including nucleophilic substitution, esterification, acylation, and

nucleophilic addition, that have been extensively used for derivatization of SAMs via solution-phase chemistry. While the effect of solvent in these reactions has not been quantified, the known solution-phase reactivity can be used to design the corresponding ion/surface reactions. For example, it has been demonstrated that similar interfacial reactions could be facilitated using reactive landing and solution-phase reactivity of lysine containing peptides with the SAM of N-hydroxysuccinimidyl ester-terminated alkylthiol on gold (NHS-SAM) [63].

Reactive landing experiments are not limited to SAM surfaces; many other reactive surfaces, including metals [30, 43, 62, 69], semiconductors [66, 70–74], polymers [75, 76], and graphite [37], could also be used in these experiments. Different types of molecules, including a variety of small molecules, peptides, proteins, lipids, oligosaccharides, and dendrimers, can be selectively immobilized on appropriate substrates using reactive landing. Finally, we note that reactive landing can also be accompanied by substantial fragmentation of the precursor ion. This phenomenon is particularly useful for modification of polymer surfaces, discussed later in this chapter.

Because reactive landing results in covalent immobilization of selected ions on surfaces while soft landing results in capture of ions or molecules via noncovalent interactions, it is usually easy to distinguish between these processes by examining the time evolution of soft landed molecules. For example, thermal desorption of the deposited species is negligible for reactive landing, but is a major decay process following soft landing. In addition, soft landed species can be removed by rinsing the surface with appropriate solvent, while reactively landed compounds are not affected by the rinsing procedure. However, it may be difficult to distinguish between adsorption involving strong hydrogen or ionic bonds and covalent binding of the molecule to the surface. It is commonly assumed that molecules that cannot be removed from the surface by rinsing are reactively landed on the substrate.

3 Instrumentation

Instrumentation utilized in soft landing experiments has been extensively reviewed and will be only briefly described here [7, 9, 10, 77]. Soft landing experiments involve four major stages: ion formation, mass selection, ion transfer to the surface, and ion deposition, and can be performed using a wide variety of ion sources and mass analyzers. For a detailed discussion of the ionization techniques in mass spectrometry, readers are referred to a number of excellent reviews [78, 79]. Electron impact ionization (EI) and chemical ionization (CI) ion sources are widely used in commercial mass spectrometers. However, these methods result in the formation of vibrationally excited ions that undergo substantial fragmentation and are limited to relatively volatile and thermally stable compounds. Soft ionization methods, such as electrospray ionization (ESI) [80], photoionization [81], matrix-assisted laser desorption ionization (MALDI) [82], fast-ion bombardment (FAB) [83], and other have been developed to produce intact molecular ions of large thermally labile molecules, e.g., biomolecules, polymers, and cluster ions. In addition, recently developed ambient surface ionization techniques facilitate formation of gas-phase

ions from a variety of liquid and solid matrices, including tissues and their subsequent purification using mass spectrometry [79]. Mass selection of the ion of interest for the soft landing experiment can be performed using different types of mass analyzers. Soft landing experiments have been performed using a linear ion trap [42], quadrupole mass filter [31, 54, 84, 85], Wien filter [65], sector mass spectrometer [55–87], Fourier transform ion cyclotron resonance mass spectrometer [39, 40, 53], and other techniques. It should be noted that several soft-landing studies have been carried out without mass selection [43, 62, 69]. Quadrupole mass filters provide a convenient way of preparing a continuous beam of m/z-selected ions using a combination of rf and dc potentials applied to the opposite pairs of rods. The ion beam exiting the mass filter is focused using ion optics and delivered to the surface. The collision energy is determined by the potential difference between the ion source and the surface and by the charge state of the ion.

4 Surface Characterization

Surface characterization is crucial for the development of robust analytical approaches for surface modification. However, physical and chemical characterization of surfaces following soft and reactive landing is challenging because of the very small amount of material (typically $<10^{13}$ molecules) deposited on the surface. Several surface characterization techniques have been utilized by different groups, including laser desorption ionization [41, 42, 54, 86], low-energy chemical sputtering [28, 86, 87], secondary ion mass spectrometry (SIMS) [39, 40, 63, 77, 86, 88, 89], surface-enhanced Raman spectroscopy (SERS) [69], X-ray photoelectron spectroscopy (XPS) [60, 76], microscopy [33, 37, 44, 45], infrared reflection absorption spectroscopy (IRRAS) [52, 59, 63, 90], temperature-programmed desorption (TPD) [45, 48, 49, 52], and a variety of biological assays [41–43]. Combination of more than one technique is often necessary to provide reliable information for understanding the underlying phenomena in reactive and soft landing.

5 Surface Modification Using Reactive Landing

5.1 Modification of Self-Assembled Monolayer Surfaces Using Small Projectile Ions

The first reactive landing experiment of intact molecular moieties was reported by Cooks and coworkers [61]. Collisions of aryl cations with an ionic surface produced by reaction of the carboxylic acid-terminated SAM surface with ammonium hydroxide resulted in decarboxylation and selective incorporation of intact aryl species such as $C_6H_5^+$, $ClC_6H_4^+$, and $BrC_6H_4^+$ into the surface. The reactive landing experiments were performed at low collision energies ($<20\,eV$) to prevent or reduce fragmentation of the reactive projectiles and to help preserve the molecular integrity of the ionic reagent. At somewhat higher collision energies, the terminal

arylhydrocarbon adsorbate formed transiently on the surface underwent benzylic C-C cleavage to yield the scattered benzylic ion. While the detailed mechanism of the reactive landing of aryl cations is unclear, it has been suggested that the reaction involves charge transfer between the projectile cation and the surface-bound carboxylate anion—a process similar to the classical Kolbe-Schmitt carboxylation reaction [91].

In another experiment from this group, formation of the Si-O bond and the terminal trimethylsilyl ether was observed following low-energy (15 eV) hyperthermal collisions of $Si(CH_3)_3^+$ ions with hydroxyl-terminated SAM (HO-SAM) surfaces [60]. The modification was confirmed using in situ chemical sputtering with 70 eV CF_3^+ ions, ex situ secondary ion mass spectrometric analysis (12 keV Ga^+ primary ion beam), and X-ray photoelectron spectroscopy by monitoring the Si (2s) peak. Comparison between the SAM surface modified by reactive landing and the synthetic SAM surfaces prepared from mixtures of the trimethylsilyl-11-mercapto-1-undecane ether and various proportions of the 11-mercapto-1-undecanol revealed that ca. 30% of the terminal hydroxyl groups were covalently modified upon hyperthermal ion bombardment. Analogous surface transformations were also achieved using $Si(CH_3)_2F^+$ and $Si(CH_3)_2C_6H_5^+$ projectile ions. In addition, surface patterning was performed by placing a nickel grid in front of the HO-SAM surface during reactive landing of $Si(CD_3)_3^+$ ions. Figure 3 shows a TOF-SIMS image obtained

Fig. 3 (*Top*) High-resolution TOF-SIMS $Si(CD_3)_3^+$ ion image from an 11-mercapto-1-undecanol self-assembled monolayer (HO-SAM) surface treated with 15 eV $Si(CD_3)_3^+$, m/z 82, ions while a Ni grid (wire space, 0.0025 in.; wire diameter, 0.0006 in.; 64% transmission) was in place in front of the surface. (*Bottom*) High-resolution TOF-SIMS spectra collected from a region modified by 15 eV $Si(CD_3)_3^+$ ions and from an unexpected control region. The primary ions in both cases were ^{69}Ga, 12 keV. Reproduced from Evans et al. [60] with permission from the American Chemical Society, copyright 2002
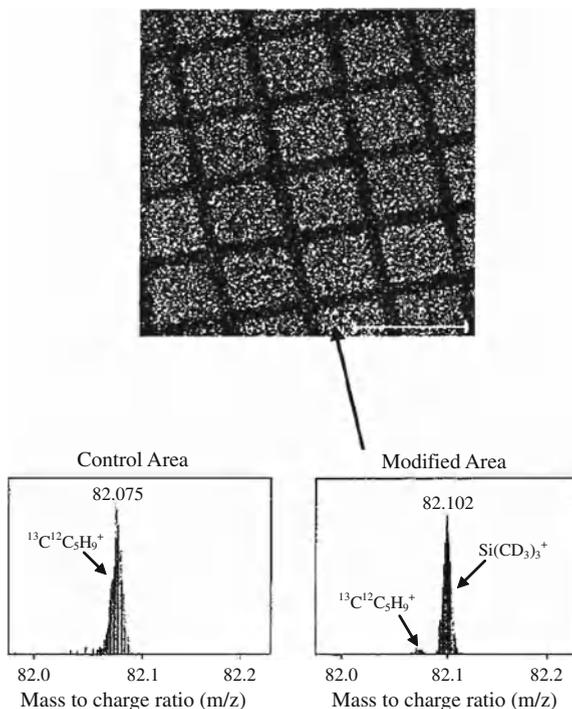
for $Si(CD_3)_3^+$ ions, nominal m/z 82, that clearly follows the grid structure, where brighter intensity regions correspond to areas exposed to the ion beam during the reactive landing experiment. High-resolution TOF-SIMS spectra obtained from the modified and unmodified regions of the surface (bottom panels) show clear separation between the ^{13}C isotope of the surface-related $C_5H_9^+$ peak at m/z 82.075 and $Si(CD_3)_3^+$ ions at m/z 82.102 obtained following exposure of the surface to the ion beam.

Other reactions between SAMs and small projectile ions reported so far include esterification and ether formation on the HO-SAM surface following collisions with $C_6H_5CO^+$ and $C_6H_5CH_2^+$ ions, respectively [59], chemical modification of FSAM surfaces using $SiCl_4^+$, $OCNCS^+$, $(CH_3)SiNCS^+$, $Si(NCO)_n^+$ ($n = 3$ and 4), $ClC(CN)_2^+$, and $CH_2Br_2^+$ projectile ions [55–58]. These studies suggest that reactive landing is an efficient noncontact method for covalent modifications of SAMs that complements current microcontact printing and other methods of nanostructure formation.

5.2 Modification of Polymer Surfaces Using Mass-Selected Ion Beams

Plasma processing is a standard industrial method for modification of polymer surfaces and thin film deposition [92–95]. Physical and chemical modification of polymers by fluorocarbon (FC) plasmas is essential for a variety of applications in biomaterials, optics, chemical sensors, and electronics [92, 94, 96–98]. For example, FC films prepared by plasma treatment have been used as diffusion barriers in drug delivery systems and for producing vascular implant devices resistant to biofouling [92, 99]. It has been shown that polyatomic ions and neutrals play a critical role in FC plasma-induced surface chemistry. Comprehensive understanding of the fundamental physical and chemical interactions that take place in plasma reactors is essential for designing successful manufacturing processes. The type and the extent of surface modification are determined by the identity and the kinetic energy distributions of ions and neutral molecules in plasmas. It follows that studies of interactions of mass- and kinetic-energy selected ions with well-defined surfaces are important for obtaining a fundamental understanding of the physical and chemical processes that control plasma-surface modification [100].

Hanley and coworkers studied chemical modification of polystyrene (PS) surfaces by low-energy (10–100 eV) SF_5^+ and $C_3F_5^+$ ions [65]. The surfaces were characterized using X-ray photoelectron spectroscopy and laser desorption ionization ion trap mass spectrometry. This study demonstrated that collisions of SF_5^+ ions with the PS surface resulted in efficient fluorination of the polymer surface. The reaction involved fragmentation of the projectile ion followed by attachment of reactive fluorine atoms to the surface. In contrast, collisions of $C_3F_5^+$ ions with PS surfaces resulted in formation of both fluorine atoms and molecular C_mF_n fragments

that subsequently reacted with the polymer surface. As a result, chemical composition of the fluorocarbon films grown on polystyrene surfaces following bombardment with hyperthermal fluorocarbon cations showed a strong dependence on the identity of the projectile ion [76]. It has been demonstrated that the distribution of different fluorocarbon functional groups on the surface and the extent of surface modification depend on the incident ion energy, the structure of the projectile ion, and the fluence. Deposition of intact projectile ions was observed for 25 eV collisions of CF_3^+ and $C_3F_5^+$ with PS surfaces, while substantial fragmentation of the incident ions was observed at collision energies above 50 eV. A higher degree of fluorination was obtained for the $C_3F_5^+$ projectile ion and attributed to more abundant formation of CF_2, $CFCF_n$, and CCF_n moieties on the surface for collisions of a larger ion with the polymer surface. At high fluences, the deposited FC species also undergo cross linking with adjacent molecules or ions to form continuous FC films. Variation of surface modification with ion flux was further utilized for preparation of chemical gradients on polymer surfaces by systematic variation of the ion current across the surface [101].

The experimental results were compared to the results of molecular dynamics simulations for collisions of the analogous hydrocarbon ions, CH_3^+ and $C_3H_5^+$, with polystyrene [75]. MD simulations demonstrate similar degrees of scattering and penetration of both projectiles into the surface for 20 eV collisions. No dissociation of the projectile ions is observed at this collision energy. In contrast, at 50 eV more than 75% of CH_3^+ ions and 65% of $C_3H_5^+$ undergo fragmentation. While a similar extent of fragmentation was observed for both ions, more efficient surface modification by $C_3H_5^+$ ions was attributed to the formation of many more reactive fragments upon dissociation of the larger projectile.

These studies demonstrate that chemical modification of surfaces using mass-selected polyatomic ions is quite selective and strongly depends on the ion structure, kinetic energy, and flux. The insights obtained by examining collisions of mass-selected hyperthermal ions with surfaces are essential for a fundamental understanding of plasma-surface interactions and for the development of new approaches for controlled modification of polymers based on their unique selectivity and the ability to prepare and transfer specific chemical functionality to the surface using mass spectrometry.

5.3 Fabrication of Nanostructured Surfaces Using Size-Selected Cluster Ions

Small nanoparticles possess unique, size-dependent chemical and physical properties that make them ideal candidates for a number of applications in catalysis and manufacturing of sensors and electronic devices. Soft landing of size-selected cluster ions onto substrates represents a new method of preparing nanostructured surfaces with lateral feature sizes in the range of 1–10 nm. Decomposition of soft-landed clusters, their diffusion and aggregation on surfaces are the major chal-

lenges in preparation of nanostructured materials using cluster deposition. While the decomposition of the cluster ion is largely determined by its initial kinetic energy, diffusion and aggregation are mainly controlled by the physical and chemical properties of surfaces. It has been demonstrated that soft landing with minimal decomposition can be achieved using low-energy deposition of clusters onto cryogenically cooled surfaces under ultrahigh vacuum conditions [84, 102]. In addition, cooling of the surface dramatically reduces the rate of diffusion and aggregation of clusters. However, most practical applications require surfaces that are stable in air or liquid at room temperature.

Vajda et al. studied thermal stability and growth of Au nanoparticles deposited on $SiO_2/Si(111)$ and $Al_2O_3/SiO_2/Si(111)$ surfaces [103]. They concluded that the final size of supported nanoparticles is determined by the initial cluster size, surface coverage and the properties of the target. The properties of the resulting materials could be further tailored using controlled heating of the substrate. An interesting phenomenon has been reported by Palmer and coworkers for deposition of size-selected Ag_{147} clusters on graphite [37]. They discovered that when the impact energy exceeds a critical value, Ag_{147} clusters are pinned to the surface, which preserves the gas-phase cluster size at elevated temperatures and prevents diffusion. The pinning threshold corresponds to the collision energy where a single carbon atom on the graphite surface is displaced from its equilibrium site, creating a reactive point defect followed by covalent binding of the cluster to the surface at the defect site. For Ag clusters incident on the graphite surface, the pinning threshold corresponds to a cluster impact energy of approximately 10 eV per Ag atom. It follows that immobilization of large clusters containing thousands of atoms requires very high impact energies that are difficult to achieve experimentally. Alternatively, point defects on the surface can be created prior to cluster deposition through irradiation of the surface with inert gas ions. Cluster pinning to the surface was also demonstrated by Kaiser et al. for collisions of positively charged antimony clusters Sb_n^+ with HOPG surfaces, but was described as "implantation" [104]. Pinning/implantation of size-selected clusters provide a unique way to prepare controlled and environmentally stable nanostructures on surfaces.

Bottom-up growth of polycrystalline thin films using specific building blocks is another novel application of size-selected cluster deposition. In contrast to the isolation of size-selected clusters on surfaces discussed in the previous paragraph, these studies take advantage of the interaction between deposited clusters and their aggregation, resulting in the formation of ordered nanoscale supramolecular architectures. Castleman and coworkers deposited mass-selected metal-carbon clusters (M_8C_{12}) on carbon-covered grids under both hard- ($\sim$2.28 keV) and soft-landing ($<$20 eV) conditions [44]. They found that high-energy collisions resulted in formation of bulk metal carbide on the surface, while soft-landed metal-carbon clusters formed an FCC structure with lattice parameter of $\sim$1.5 nm, corresponding to the cluster assembly. Müllen and coworkers soft landed large graphene ions generated by matrix-assisted laser desorption/ionization (MALDI) onto HOPG surfaces [47]. Subsequent STM characterization of the surfaces revealed that the graphene molecules packed "edge-on" at the surface, forming ordered semiconducting architectures.

The structure of the films formed by soft landing was remarkably different from the assembly of graphene films prepared by vacuum sublimation [105] or solution processing [106], where the graphene discs adsorb "face-on" on the basal plane of the HOPG, forming hexagonally packed layers. The difference was attributed to the difference in affinities of charged and neutral graphene molecules to the graphite surface that controls the assembly.

Kappes and coworkers demonstrated that soft landing of low-energy (typically 6 ± 0.5 eV) mass-selected C_{58}^{+} ions onto the HOPG surface resulted in the formation of a new solid material through aggregation of C_{58} molecules deposited on the target with the incoming projectile ions [107]. Growth of the solid C_{58} phase proceeds according to the cluster-aggregation-based Volmer-Weber scenario, where initially ramified 2D fractal islands transform into 3D dendritic mounds at higher coverage. The physical properties of the resulting C_{58} films are very different from those of the C_{60} solid phase. Figure 4 shows AFM images of C_{58} and C_{60} layers deposited onto HOPG under identical conditions (kinetic energy, surface tempera-
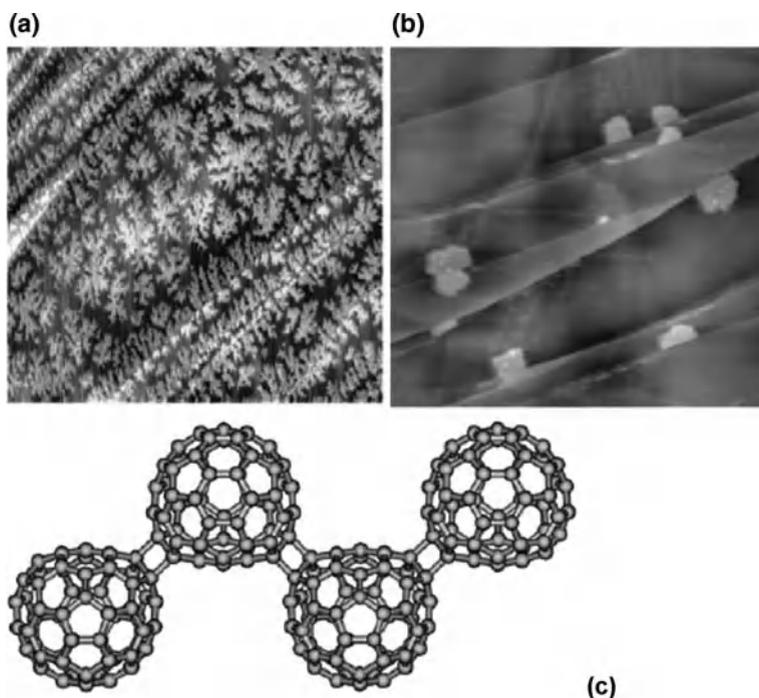
Fig. 4 (**a**) and (**b**) AFM images (image size $3 \times 3 \, \mu m^2$) contrast the topography of the C_{58} and C_{60} films, respectively. Both images were taken using noncontact imaging mode after deposition of the same ion amounts of 1.2×10^{13} (9.2×10^{13} ions cm^{-2}) at 300 K and the kinetic energy of 1 eV. The height profile derived from the tip scans reveals a mean thickness of C_{60} layers of around 1 nm and a significantly lower height of 0.8 nm for C_{58} islands; (**c**) possible four-oligomer structure resulting from covalently linking C_{58} (1) into a chain, as predicted by BP86/SVP calculations. Reproduced from Bottcher et al. [107] with permission from the Royal Society of Chemistry, copyright 2005

ture, ionic flux, and coverage). While C_{60} forms rather compact islands terminated by smooth rims always associated with HOPG step edges, C_{58} arranges in fractal-like islands that decorate not only step edges but also flat terraces. In addition, thermal desorption spectra reveal that C_{58} films are thermally more stable than C_{60} films. It follows that the cluster-cluster and/or cluster-surface interactions in C_{58} films are much stronger than the corresponding interactions in the C_{60} layers deposited onto HOPG surfaces. DFT calculations suggest that C_{58} undergoes efficient aggregation on the surface. The aggregation is governed by reactive sites comprising adjacent pentagons or heptagons on individual cages that are absent in C_{60} cages. The resulting covalent cage-cage bonds shown in Fig. 4 are responsible for the high thermal stability of C_{58} films.

In conclusion, soft landing/reactive landing of size-selected cluster ions onto surfaces provides a unique way to generate novel nanoparticles, fabricate nanostructures, and prepare new solid materials using mass-selected cluster ion beams. The properties of the films can be precisely controlled by varying the kinetic energy of the cluster ions during deposition.

5.4 Preparation of Biomaterials Using Reactive Landing

The soft ionization methods discussed earlier can be used for introducing a variety of biological molecules into the gas phase as singly or multiply charged ions. Deposition of mass-selected ions of biological molecules onto substrates enables very selective preparation of novel biomaterials using soft or reactive landing. In addition, soft landing can be used for separation and purification of biomolecules from complex mixtures for subsequent analysis. It has been demonstrated that ionized complex biomolecules such as DNA [53], proteins [41–43], and even whole viruses [54], can be recovered after soft landing with retention of biological activity.

In a study by the Cooks group, protein arrays were prepared by soft landing of selected charge states of multiply protonated proteins, including cytochrome c, lysozyme, insulin, and apomyoglobin at different positions on a gold surface [42]. Soft-landed proteins were subsequently rinsed from the surface and analyzed by electrospray ionization mass spectrometry. Mass spectra obtained from rinse solutions were identical to the spectra of authentic compounds, suggesting that ions were deposited nondestructively on the surface. In addition, retention of bioactivity was confirmed for trypsin, lysozyme, and different kinases following the soft landing process, demonstrating the utility of soft landing for selective preparation of protein arrays. The soft landing efficiency was independent of the selected charge state of the protein. However, the collision energy and the properties of the substrate had a profound effect on protein soft landing, with fragmentation (crash-landing) observed at collision energies in excess of 10 eV/charge.

Tureček and coworkers conducted preparative soft and reactive landings of singly charged polyatomic organic ions and multiply charged protein ions on plasma-treated metal surfaces [43, 62]. They found that a significant fraction of a

fluorescence-labeled biotin conjugate soft landed at 40–50 eV kinetic energy could not be washed from the surface, suggesting stable immobilization of these species by soft landing. In addition, retention of biological activity was reported for immobilized trypsin and streptavidin ions soft-landed at nominal kinetic energies of 130–200 eV. This group further demonstrated the utility of reactive landing for the preparation of biomedically useful coatings using immobilization of hyaluronan—a major component of the extracellular matrix—on stainless steel surfaces [108]. Hyaluronan is a polysaccharide composed of hundreds to thousands of repeating disaccharide units, as shown in Fig. 5. Immobilization of hyaluronan on hard stainless steel surfaces is of great importance for preparation of novel coatings for improved biocompatibility of endovascular stents. Substrates prepared by soft landing of hyaluronan were extensively washed in polar solvents and submitted for platelet adhesion studies. Platelet adhesion was quantified using scanning electron microscopy (SEM). SEM images, shown in Fig. 5, indicate efficient platelet adhe-
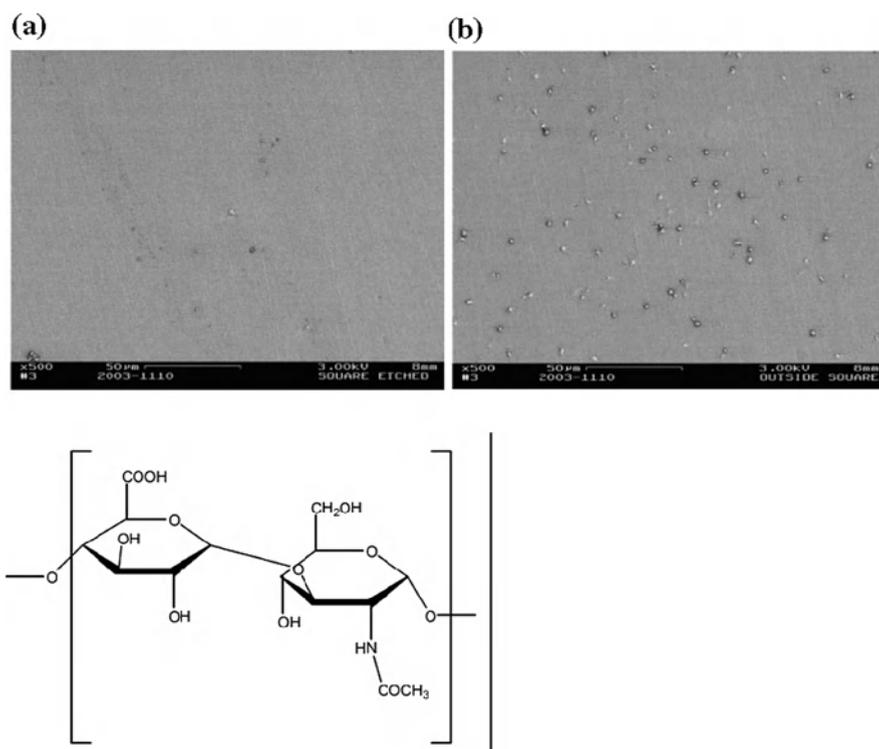
Fig. 5 *Top*: SEM image of the sample surface that was washed and exposed to the platelet adhesion study (**a**) inside the spot of the deposited hyaluronan; (**b**) outside the hyaluronan spot on the blank surface (no hyaluronan deposited). *Bottom*: Disaccharide structure of hyaluronan. This disaccharide unit is repeated (250–25,000 times) resulting in hyaluronan chains with molecular mass up to several million Daltons. Reproduced from Volný et al. [108] with permission from Wiley InterScience, copyright 2007

sion in the uncoated area and almost no adhesion in the area modified with hyaluronan, suggesting that the immobilized hyaluronan retained its protective properties against blood platelet activation. In these studies, surfaces were characterized using bioassays, fluorescence spectroscopy, scanning electron microscopy (SEM), and atomic force microscopy (AFM). Although these methods enable the detection of the soft-landed material and provide valuable information on its biological activity, they cannot reveal the nature of bonding to the surface and are not sensitive to possible chemical transformations of the immobilized molecules. In a subsequent soft landing study of polyatomic cations, including organic dyes and nucleosides, on plasma-treated silver substrates, surface-enhanced Raman Scattering (SERS) was utilized for structural characterization of immobilized molecules [69]. SERS spectra from submonolayer amounts of soft-landed compounds were sufficiently intense and reproducible to allow identification of Raman active vibrational modes for structure assignment. SERS characterization demonstrated that tethered molecules are microsolvated on the surface and bound via ion pairing or π-complexation to silver atoms and ions in the surface oxide layer. Comparison of spectra obtained following reactive landing and solution-phase chemistry indicates that the molecules survive soft landing without significant chemical damage.

Recently, Wang et al. reported the first study of covalent immobilization of peptides on SAM surfaces using reactive landing of RGD-containing peptides [63, 109]. The RGD (arginine-glycine-aspartate) motif is by far the most effective and most frequently employed peptide sequence for stimulated cell adhesion on synthetic surfaces. It has been demonstrated that RGD-containing peptides promote cell adhesion when immobilized on surfaces, and inhibit it when presented to cells in solution [110], suggesting that covalent linking of RGD peptides to substrates is essential for strong cell adhesion. Soft landing of a doubly protonated cyclic pentapeptide c(-RGDfK-) onto a SAM surface terminated with an active N-hydroxysuccinimidyl ester group (NHS-SAM) resulted in covalent linking of the peptide to the SAM surface through formation of an amide bond between the surface and the amino group of the lysine side chain, as shown in Fig. 6. Figure 7 shows the results of the TOF-SIMS and IRRAS characterization of the modified surface. TOF-SIMS spectrum of the NHS-SAM following soft landing of c(-RGDfK-) contains abundant singly protonated peptide ion, $[M + H]^+$, and a characteristic doublet of the protonated covalent adduct (CA). The CA doublet contains the peak at m/z 804.4, corresponding to the combined molecular weight of the protonated peptide (604.3) and the thiol (315.2), without the NHS endgroup (115.1) and its analog at m/z 802.4 that has an additional double bond. The $[M + H]^+$ peak corresponds to loosely bound peptide molecules on the surface and can be readily removed by ultrasonic washing of the NHS-SAM surface in methanol. In contrast, the $[CA+H]^+$ doublet is not affected by rinsing, suggesting that it is a signature of the covalently bound peptide.

Covalent binding is further confirmed using IRRAS characterization. Amide I and amide II bands at ca. 1675 and 1535 cm^{-1}, respectively, observed in IRRAS spectra following soft landing of c(-RGDfK-) (Fig. 7b), originate both from peptide bonds of soft-landed peptides and from the newly formed amide bond between the
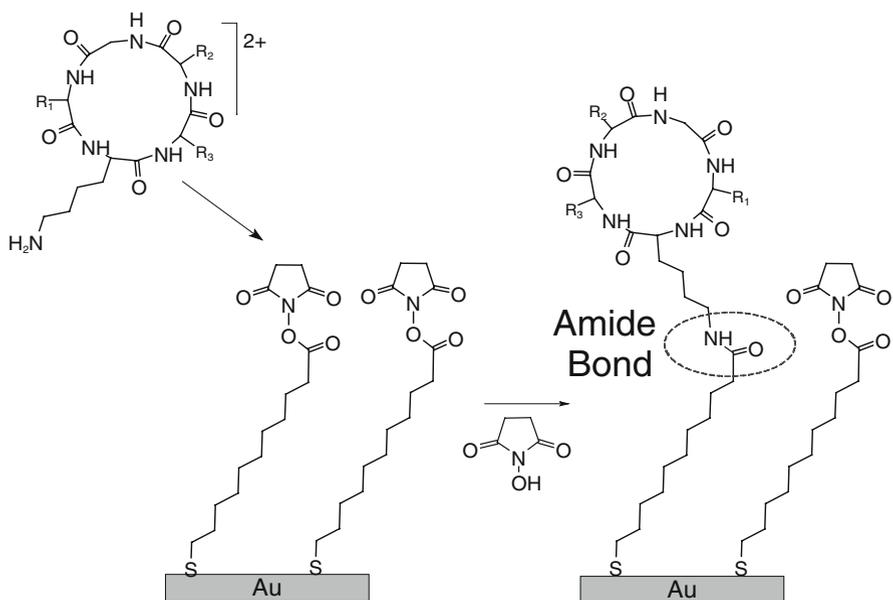
Fig. 6 Reactive landing of doubly protonated c(-RGDfK-) on NHS-SAM. Reproduced from Wang et al. [63] with permission from the American Chemical Society, copyright 2007

peptide and the SAM. Rinsing of the surface removes most of the loosely bound molecules from the surface, resulting in depletion of the amide bands in the IRRAS spectra. The retention of the residual amide bands along with the suppression of the bands characteristic of the NHS group (the asymmetric stretching of the NHS
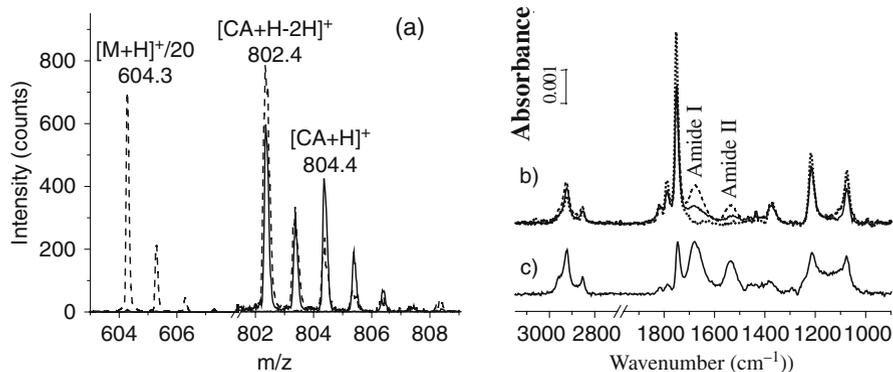
Fig. 7 (**a**) and (**b**) TOF-SIMS and IRRAS spectra following 4 h soft landing of 1.8×10^{13} doubly protonated c(-RGDfK-) molecules on NHS-SAM. *Dotted line*: blank NHS-SAM; *dashed line*: NHS-SAM following soft landing; *solid line*: soft landed sample with extensive rinsing in methanol. (**c**) IRRAS spectrum of the NHS-SAM following 2 h solution phase reaction. Reproduced from Wang et al. [63] with permission from the American Chemical Society, copyright 2007

carbonyls at $1751\,cm^{-1}$, the asymmetric CNC stretching of the NHS at $1217\,cm^{-1}$, and the NCO stretching of the NHS at $1074\,cm^{-1}$) indicate that a fraction of the peptides are covalently linked to the SAM surface. Comparison of the IR intensities of amide I and amide II bands of the peptide, immobilized using solution phase chemistry (Fig. 7c) and reactive landing (Fig. 7b), demonstrated that similar local coverage of 60% of a monolayer was obtained using 4 h soft landing and 2 h reaction in solution.

6 Factors that Affect the Efficiency of Reactive Landing

Understanding the factors that determine the efficiency of ion-surface reactions is crucial for the development of robust surface modification approaches. The following section discusses the effects of the structure, property, and collision energy of the projectile ion, as well as the physical and chemical properties of the surfaces, on the efficiency of reactive landing.

6.1 Structure and Properties of the Projectile Ion

Because reactive landing involves both trapping of ions or molecules on surfaces and an interfacial chemical reaction, its efficiency is determined by a combination of the soft landing efficiency and the reaction yield. Factors affecting trapping of mass-selected ions on surfaces have been recently reviewed [7, 9, 10, 111]. For a given kinetic energy of the projectile ion the reaction yield depends on the structure and property of the ion and the terminal groups on the reactive surface. Cooks and coworkers examined the efficiency of silylation of the HO-SAM surface using a set of silylium cations listed in the first column of Table 1 [112]. The efficiency of silylation reaction was determined using chemical sputtering. The ratios of chemically sputtered ions, representing the efficiency of reactive landing of different projectiles, are affected by the partial charge on the central silicon atom that is altered by substitution. The results demonstrate that the reactivity increases with an increase in the partial charge on the silicon atom. For example, because carbon is more electronegative than silicon, $Si(CH_3)_3^+$ bears a strong positive charge on Si and readily reacts with the HO-SAM surface. In contrast, $SiCl_3^+$, which has a lower positive charge on its silicon atom because of the π-orbital overlap between the Si and Cl atoms, does not react with the HO-SAM. Some of the discrepancies between the atomic charge density on Si and the reactivity were attributed to differences in sputtering efficiencies of different chemical moieties on the surface. Because the reaction proceeds via direct electrophilic attack of the cationic projectiles on the surface hydroxyl functionalities, the electrophilicity of the reagent ion determines the rate of the silylation reaction. This reaction mechanism rationalizes the observed increase in reactivity of silyl cations with the increase of the positive charge on Si.

Table 1 Effect of chemical structure and local atomic charge on silylation efficiency

Reagent ion[a]	Atomic charge density on silicon (HF/6–31 G)[b]	Ratio of chemically sputtered ions (%)[c]
$SiCl_3^+$	0.953	0
$SiBr_3^+$	0.977	0
$Si(CH_3)_2Cl^+$	1.088	0
$Si(CH_3)_2Br^+$	1.091	0
$Si(CH_3)_2CN^+$	1.129	9
$Si(CH_3)_2(C_6H_5)^+$	1.133	46
$Si(CH_3)_3^+$	1.152	84
$Si(C_2H_5)_3^+$	1.164	85
$Si(CH_3)_2NCO^+$	1.182	42
$Si(CH_3)_2NC_2H_6^+$	1.182	75
$Si(CH_3)_2CH_2Cl^+$	1.183	53
$Si(CH_3)_2F^+$	1.401	72
$Si(OCH_3)_3^+$	1.603	56

[a] Silylium ions used in attempted ion–surface modification experiments.
[b] Local atomic charge associated with silicon atom of the reagent ion as calculated by Hartree–Fock ab initio theory with a basis set of 6–31 G(d, p).
[c] Percentage of chemically sputtered ions observed in the 70 eV scattered ion mass spectra of CF3C that result from a modification reaction.
Reproduced from Wade et al. [112] with permission from Wiley InterScience, copyright 2002.

Other chemical factors such as steric hindrance and the stability of the projectile ion also affect the reactivity.

The structure of the projectile ion has also a strong effect on the reactive landing of peptide ions on SAM surfaces. Laskin and coworkers examined the reactivity of the N-terminal α-amino group and the ε-amino group of the lysine side chain during the reactive landing of peptide ions onto NHS-SAM surfaces [109]. They compared the reaction efficiency of the cyclic c(-RGDfK-) peptide and two linear RGD-containing peptides with and without the lysine residues GRGDSPK and RGDGG. In contrast with the cyclic peptide that can be bound to the NHS-SAM only via the lysine side chain, both linear peptides contain a primary amino group at the N-terminus that can also participate in covalent binding of the peptide to the SAM surface. However, TOF-SIMS and IRRAS characterization indicate very inefficient binding of the peptides through the N-terminal amino group and efficient linking through the ε-amino group of the lysine side chain. Since primary amines react with NHS esters by nucleophilic attack, resulting in formation of an amide bond between the amine and the ester accompanied by the release of the N-hydroxysuccinimide (NHS), the enhanced reactivity of the ε-amino group in the gas phase could be attributed to the stronger nucleophilicity of this group as compared to the N-terminal amino group.

Hanley and coworkers studied the effect of the primary structure of the projectile ion on reactive landing in modification of polystyrene surfaces using low energy CF_3^+ and $C_3F_5^+$ ion beams [75]. Figure 8 displays the C(1s) region of the XPS spectrum of unmodified PS (bottom trace) and PS exposed to 25 eV $C_3F_5^+$ modified

Fig. 8 C(1s) region of X-ray photoelectron spectra of native polystyrene (PS) thin film (*bottom*), 25 eV CF_3^+ modified PS, 25 eV $C_3F_5^+$ modified PS, 50 eV CF_3^+ modified PS, and 50 eV $C_3F_5^+$ modified PS (*top*). Components of C(1s) peak are labeled, with detected carbon atom *underlined*. All ion fluences are equivalent to 1.5×10^{16} Farad atoms/cm^2. Reproduced from Wijesundara et al. [75] with permission from the National Academy of Sciences, copyright 2000
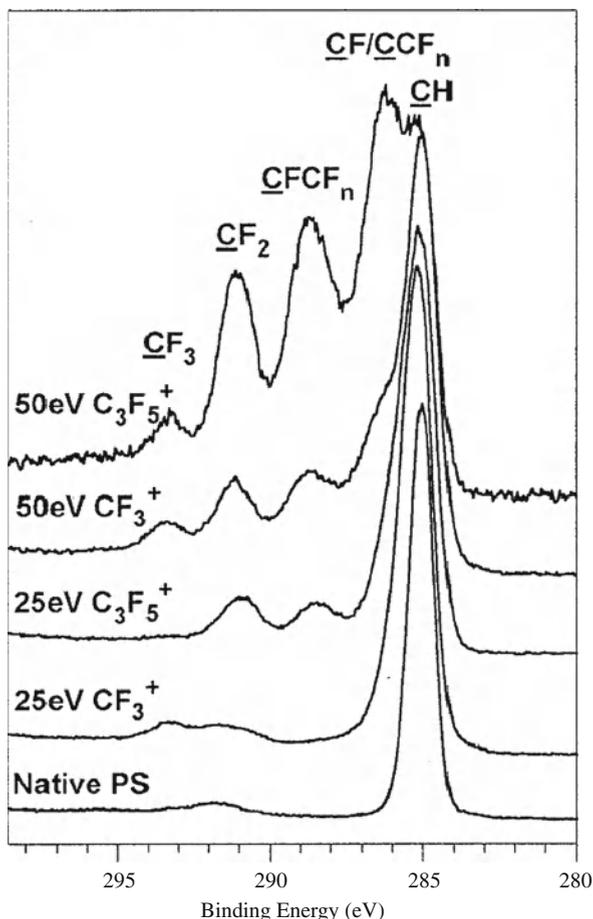
PS, 50 eV CF_3^+ modified PS, and 50 eV $C_3F_5^+$ (top). The results reveal more efficient growth of polymer films using $C_3F_5^+$ ions as compared to CF_3^+ projectiles at a similar energy/atom of ≈ 6 eV/atom or similar total kinetic energies of 25 and 50 eV. Molecular dynamics simulations utilized in these studies provide important insight into various physical and chemical phenomena associated with the polymer film growth following ion-surface collisions. Specifically, MD simulations were used to explore energy transfer in collisions and the competition between scattering, ion penetration, deposition, and dissociation of projectile ions upon impact. The effect of the size and complexity of the projectile ion on the extent of surface modification determined experimentally was attributed to differences in dissociation, reactivity, sticking probabilities, and energy transfer of different precursor ions to the surface [75, 76].

6.2 Properties of the SAM Surface

The reactivity of the target surface is also a crucial factor that determines the rate of the interfacial reaction and hence the efficiency of reactive landing. In particular, when SAM is used as the soft-landing target, the presence of active functional groups on the surface plays a key role in the outcome of the reaction. The importance of surface activation in reactive landing of aryl cations on carboxylic acid-terminated SAM surfaces was mentioned earlier. Specifically, no reaction was observed following collisions of $C_6H_5^+$ ions with a COOH-SAM surface, while collisions with the ionic ammonium salt-terminated surface formed by treatment of the COOH-SAM with ammonium hydroxide resulted in decarboxylation of the terminal surface groups and subsequent C-C bond formation between the projectile and the surface [61].

The efficiency of covalent immobilization of peptides on SAM surfaces using reactive landing also depends on the properties of the surface. Specifically, reactive landing of peptide ions on SAM surfaces requires the presence of a labile NHS ester terminal group on the surface, while no reaction is observed between peptides and the COOH-SAM surface. Similar trends were reported for solution-phase reactivity of these substrates. It has been demonstrated that activation of the COOH-SAM surface with the NHS group expedites the reaction between SAMs and amine-containing molecules in solution [113]. The improved reactivity of SAMs with an active NHS group on the surface is attributed to the properties of the molecule that is detached from the surface following the reaction, the leaving group. The NHS group is a good leaving group because it is much more labile than water that is formed when an amino group reacts with a COOH-SAM surface. The NHS functionality is often used to introduce a variety of amine-terminated molecules to SAMs under mild conditions. The reaction efficiency could be further improved using carboxylic acid-terminated SAMs activated with pentafluorophenyl esters [114] or SAMs terminated with the interchain anhydride [115].

In addition to formation of amide bond and alkane synthesis from organic carboxylates that have been demonstrated using collisions of gas-phase projectile ions with SAMs, a number of other classes of interfacial reactions can be explored in future studies. Such experiments can be designed based on the known solution phase reactivity of SAMs. Examples of possible reactions include modification of SAMs terminated with maleimide functional groups by reaction with biologically active thiols, which occurs efficiently in solution [116]; or cross-metathesis reaction for attachment of acrylamide, acrylic acid, and methyl acrylate to vinyl-terminated SAMs [117].

Physical properties of the SAM surface, including the organization of the chains in the monolayer, the density and orientation of the functional groups on the surface, lateral steric effects, and the partitioning of the free reactants at the interface, can also influence its reactivity [68]. While highly organized monolayers are very useful for a variety of applications, the limited reactivity resulting from the confinement effect may prevent efficient chemical surface modifications. It has been shown that compositionally and structurally defined yet disordered systems possess

clear advantages for chemical modification of SAMs [118]. Mixed SAM surfaces or SAM surfaces prepared using short-chain thiols that do not form a well-organized monolayer might provide a favorable environment for the approaching reagent and therefore improve the reaction yield in reactive landing.

6.3 Kinetic Energy of the Projectile Ion

The kinetic energy of the projectile ion in soft landing experiments is determined by the potential difference between the ion source and the surface that can be readily varied from several millivolts to several kilovolts. Cooks and coworkers examined the effect of the kinetic energy of the projectile ion on the esterification reaction following collisions of benzoyl cations with the HO-SAM surface [59, 112]. Figure 9a shows the extent of the reaction as a function of the collision energy of the reagent ion. The reaction yield is negligibly small below 7 eV, increases rapidly at collision energies between 7 and 15 eV, and then steadily decreases with increasing collision energy until it no longer occurs above 60 eV. Because esterification involving benzoyl cations is an endothermic reaction, the observed threshold behavior at low collision energies indicates that the kinetic energy of the projectile ion provides the energy necessary to overcome the reaction barrier. The decrease of the reactive landing efficiency at high collision energies is attributed to fragmentation of the precursor ion, which becomes more prominent as the collision energy increases.
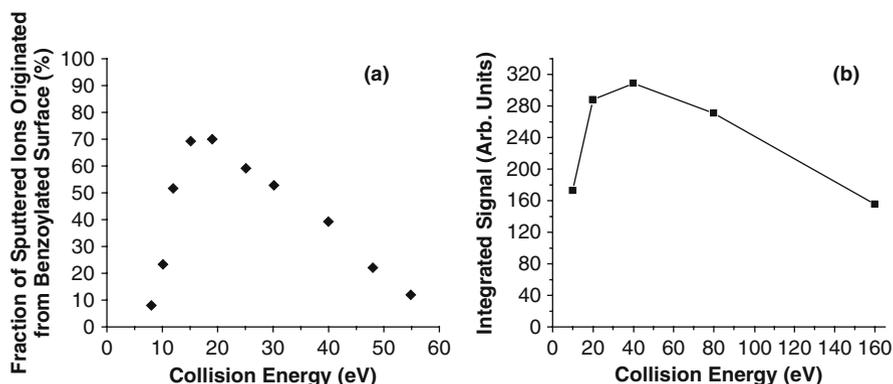

Fig. 9 (**a**) Scattered-ion mass spectra were recorded upon 70 eV collisions of CF_3^+ with several samples of the HO-SAM surface. These surfaces were modified for 2 h with the benzoyl cation at varying collision energies over the range 5–60 eV. The reaction efficiency was measured and plotted using the fraction of sputtered ions originating from the benzoylation modification vs. the total number of sputtered ions originating from the modified and unmodified portions of the HO-SAM. Reproduced from Wade et al. [59] with permission from the American Chemical Society, copyright 2002. (**b**) Total CA signal as a function of kinetic energy in the range of 10–160 eV following soft landing of 2×10^{12} doubly protonated c(-RGDfK-) onto NHS-SAM surface. Reproduced from Wang et al. [109] with permission from the Royal Society of Chemistry, copyright 2008

Similar dependence of the reaction yield on collision energy was observed for silylation of HO-SAM surface using $Si(CD_3)_3^+$ ions [112] and for immobilization of peptide ions on surfaces [109]. Figure 9b shows the dependence of the reactive landing efficiency on the kinetic energy of the projectile ions for collisions of the doubly protonated c(-RGDfK-) with the NHS-SAM surface. For this system the initial increase of the reactive landing efficiency at collision energies below 20 eV is followed by a plateau region from 20 to 80 eV and a gradual decrease at collision energies above 100 eV. At low kinetic energies (<20 eV) the reactive landing efficiency is determined by the ability of the system to overcome the reaction barrier, resulting in the observed increase in the reaction yield with collision energy. Because no crash-landing was observed for collisions of peptide ions with SAM surfaces, the decrease in the reaction efficiency observed at higher collision energies was attributed to the corresponding decrease in the soft landing efficiency discussed in detail elsewhere [40]. In that study it was demonstrated that a slow decrease in the fraction of peptide ions trapped on the surface with an increase in collision energy between 0 and 100 eV is followed by a fairly sharp decrease observed at energies above 100 eV because of the competition between the soft landing and scattering of ions at higher collision energies [8, 119].

While the qualitative trend of the dependence of the reactive landing efficiency on collision energy shown in Fig. 9 is similar for different projectiles, the collision energy at which the reaction yield is maximized may vary significantly for different projectiles and surfaces. For example, Turecek and coworkers showed that singly charged biotin conjugate can be tethered to plasma-treated metal surfaces by soft landing at 40–50 eV [62], while partial immobilization of multiply charged trypsin ions on the same surface requires nominal kinetic energies of 130–200 eV [43]. The above examples indicate that the optimal kinetic energy for reactive landing depends on the type of reaction, the size of the projectile ion and the properties of the surface. In addition, at higher collision energies, reactive landing of small polyatomic molecular ions experience severe crash-landing [65, 76]; while for larger molecules like peptides or proteins, crash-landing does not occur to a significant extent [40]. The difference in behavior between peptide ions and small organic ions is attributed to the large number of vibrational degrees of freedom that can effectively store the excess internal energy acquired by the peptide ion in the collision process, accompanied by efficient energy transfer into the substrate prior to dissociation. Because crash-landing alters the identity of the gas-phase reagent, it is reasonable to assume that intact immobilization of small projectile ions is favored only at low collision energies.

6.4 Charge State of the Projectile Ion

Ionization of biomolecules, large clusters, organometallic complexes, and polymers commonly results in the formation of multiply charged ions with a distribution of charge states. While ion-surface chemistry may be affected by the charge state of

the precursor ion, little is known about the effect of the charge state of the projectile ion on the efficiency of soft and reactive landing. Charging of functional groups may enhance or inhibit their reactivity. For example, the aminolysis of the NHS-SAM proceeds via nucleophilic attack of the free primary amines. Because the nucleophilic character of the amine is readily quenched by protonation, the location of the ionizing proton can have a pronounced effect on gas-phase reactivity of the amine containing biomolecular ions. The effect of the charge state of the ion on the efficiency of reactive landing was explored by comparing reaction yields following soft landing of singly and doubly protonated c(-RGDfK-) molecules onto the NHS-SAM surface [109]. Because of its cyclic structure, this peptide can be linked to the surface only via the ε-amino group of the lysine side chain that is most likely neutral for the $[M + H]^+$ ion and is protonated for $[M + 2H]^{2+}$ ion of c(-RGDfK-). It follows that while singly protonated c(-RGDfK-) has a free primary amino group necessary for reaction, the lysine side chain of the doubly protonated c(-RGDfK-) is not available for reaction with the NHS-SAM surface. However, similar reactive landing efficiencies obtained for the $[M + H]^+$ and $[M + 2H]^{2+}$ ions of c(-RGDfK-) indicate that doubly protonated c(-RGDfK-) is completely or partially neutralized upon collision with the NHS-SAM surface, and the reaction most likely involves the neutral ε-amino group of the lysine residue.

Charge reduction and retention—one of the important fundamental issues in soft landing—may be an important factor in reactive landing. Collisions of ions with clean metal targets usually result in neutralization of more than 99% of projectile ions. Organic thin films on metal substrates substantially reduce neutralization [9]. Cooks and coworkers have demonstrated that intact polyatomic ions could be trapped in the FSAM for many days [28]. Recently, Laskin and coworkers conducted the first study of the kinetics of desorption and charge reduction following soft landing of the doubly protonated Gramicidin S (GS) onto the FSAM surface [88]. The results indicate that a significant fraction of the doubly protonated projectile ions retain their charge on the FSAM surface, while some ions undergo instantaneous loss of one or two protons upon collision. Charge retention may not be important for reactive hydrophilic surfaces because the neutralization efficiency shows a rapid increase in the order FSAM<HSAM<COOH-SAM [89]. However, it could play an important role in modification of inert FSAM or HSAM surfaces by multiply charged projectiles.

The mechanism of neutralization of complex ions is currently unclear. Tureček and coworkers proposed that ion discharge on plasma-treated stainless steel surfaces is mediated by proton transfer from the projectile ions to the metal oxide layer and coupled with a redox reaction that is compensated by electron transfer between the oxide surface and the conduction band of the bulk metal [43, 108]. However, for SAM surfaces, electron transfer through the well-assembled organic layer is a fairly slow process. Although it has been suggested that charge could be carried away by hydrophilic terminal groups [4], the mechanism of this process has not been explored in detail.

6.5 Comparison with Surface Modification Using Reactions in Solution

Modification of SAM surfaces using solution phase reaction is largely mediated by the solvent. Various factors, including concentration of the reagent, presence of the catalyst, polarity, viscosity and solubility of the solvent, pH, and temperature of the solution, determine the reaction efficiency. In contrast, because the reactive landing experiment is carried out in the absence of solvent, the reactivity between well-organized monolayer surfaces and the gas phase reactant could be rather different. Wang et al. conducted the solution-phase reaction between GRGDSPK and the NHS-SAM surface by immersing the surface in a 1.5 mM solution of GRGDSPK in the ammonium bicarbonate (NH_4HCO_3) buffer (pH = 8) [109]. IRRAS characterization showed that only a small amount of GRGDSPK was grafted onto the surface after 15 h of reaction in solution (Fig. 10a), while similar intensity amide bands were obtained following soft landing of ca. 1.8×10^{13} of doubly protonated GRGDSPK (4 h) on the NHS-SAM surface (Fig. 10b). Because the soft-landed peptide spot is only 5 mm in diameter, while solution phase reaction generates a uniform peptide layer that spreads over the whole IR interrogated area of 10 mm, the local coverage of GRGDSPK on the NHS-SAM surface obtained using 4 h reactive landing is actually ca. 4 times higher than the local coverage obtained after 15 h of reaction in solution. Comparison between the reactivity of the GRGDSPK peptide with the NHS-SAM surface in solution and in the gas phase suggests that the relatively high kinetic energy of projectile ions utilized in reactive landing experiments can promote reactions on surfaces that are characterized by very slow kinetics in solution. It is reasonable to assume that ion-surface collision results in a local heating of the surface that facilitates the reactivity.

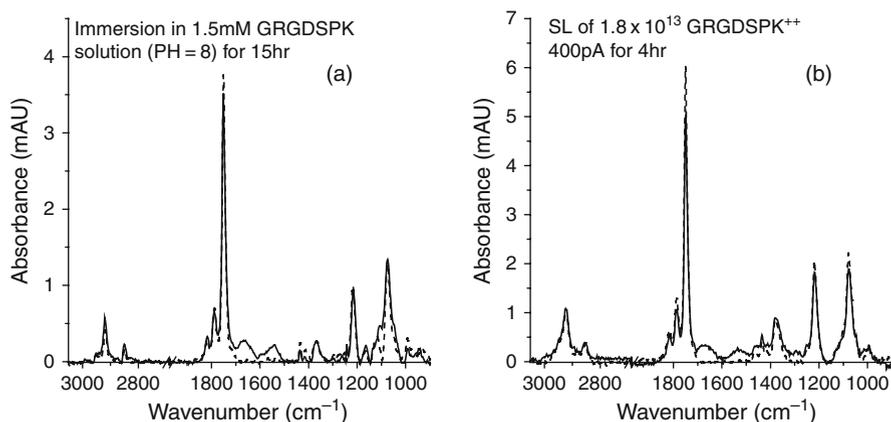**Fig. 10** IRRAS spectra of NHS-SAM (*black trace* in **a** and **b**); NHS-SAM following 15 h reaction in 1.5 mM GRGDSPK solution (pH = 8) (**a**, *red*); and NHS-SAM following soft landing of 1.8×10^{13} doubly protonated GRGDSPK and subsequent rinsing (**b**, *blue*). Reproduced from Wang et al. [109] with permission from the Royal Society of Chemistry, copyright 2008

In summary, reactive landing of mass- and energy-selected polyatomic ions on surfaces presents an important alternative to currently existing surface modification techniques. The unique advantage of reactive landing as compared to surface modification using solution-phase reactivity is the ability to control the kinetic energy of projectile ions that is essential for overcoming the reaction barrier. In addition, reactive landing introduces exceptional selectivity and specificity into the surface preparation step by eliminating the effect of solvent and sample contamination on the quality of the film, and enables surface patterning by shaping and precise positioning of the ion beam on the surface.

7 Future Directions

A variety of examples presented in this chapter indicate that reactive landing is a unique approach for highly selective surface modification that can be utilized for preparation of novel biological, polymer, and nanostructured materials. While the general mechanism of reactive landing is not well understood, many factors that determine the efficiency of this process have been identified. These include the composition and structure of the projectile ion, its kinetic energy and charge state, and the physical and chemical properties of the surface. Fundamental understanding of physical and chemical processes occurring in reactive landing experiments is essential for the development of practical manufacturing approaches involving this technique. Future studies will combine detailed experimental investigations with reliable and accurate theoretical modeling of the underlying phenomena.

Improved physical and chemical characterization of surfaces following soft and reactive landing is crucial for the development of robust analytical methods for biomaterial applications. Current surface characterization techniques including AFM/STM/EFM, TOF-SIMS, IRRAS, XPS, HREEL, and SERS provide limited structural information and are generally not designed for in situ and real-time characterization of modified surfaces. Future development in our laboratory will focus on coupling of several surface techniques with the soft landing apparatus for obtaining precise control and detailed understanding of the surface modification process.

Several technical challenges important for practical applications must be also addressed. While keV ion beams are commonly used in microfabrication, bombardment of substrates with high-energy ions results in local damage of the surface. In contrast, hyperthermal ion beams are ideally suited for controlled chemical modification of a wide variety of surfaces without any significant damage. However, nanofabrication requires special ion optics that could focus the ion beam down to submicron size—a nontrivial task for hyperthermal ion beams. In addition, while EI and CI are capable of producing intense ion beams, ion currents that can be reliably obtained using ESI are 2–3 orders of magnitude lower than ion currents utilized in micro- and nanofabrication. It follows that significant instrument development is required for possible high-throughput applications of reactive landing of electrosprayed ions.

Surface modification using reactive landing takes advantage of the selectivity and specificity inherent to mass spectrometry and ion beam chemistry. These features could contribute to scientific and technological breakthroughs in different research areas, including biochemistry, catalysis, optics, electronics, and nanotechnology. Patterning of the incident ion beam on surfaces is important for fabrication of novel devices with controlled properties. For example, hyperthermal deposition of fullerenes (C_{60}) results in formation of nanostructured thin films, in which the components of the film retain the electronic structure of the deposited molecules; nanostructures formed by deposition of metal cluster ions can be used to prepare novel cluster catalyst materials stabilized against sintering. However, efficient reactive landing with retention of the desired properties of the projectile is a major challenge that requires careful design and selection of the reactive projectile-surface pair. Finally, we note that a variety of different classes of interfacial reactions can be induced using ion-surface collisions, which makes reactive landing an attractive tool for studying reactions at interfaces in the absence of solvent.

References

1. Benninghoven A, Rudenauer FG and Werner HW, Secondary Ion Mass Spectrometry: Basic Concepts, Instrumental Aspects, Applications and Trends, Chemical Analysis; Wiley: New York, 1987; Vol. 86.
2. Pachuta SJ and Cooks RG, Chem. Rev. **87**, 647, 1987.
3. Vickerman JC and Briggs D, ToF-SIMS: Surface Analysis by Mass Spectrometry; IM Publications and SurfaceSpectra Limited: West Sussex and Manchester, UK, 2001.
4. Pacholski ML and Winograd N, Chem. Rev. **99**, 2977, 1999.
5. Rabalais JW, Principles and Applications of Ion Scattering Spectroscopy: Surface Chemical and Structural Analysis; Wiley InterScience: Hoboken, NJ, 2002.
6. Dongre AR, Somogyi A and Wysocki VH, J. Mass Spectrom. **31**, 339, 1996.
7. Grill V, Shen J, Evans C and Cooks RG, Rev. Sci. Instrum. **72**, 3149, 2001.
8. Laskin J and Futrell JH, Mass Spectrom. Rev. **22**, 158, 2003.
9. Gologan B, Green JR, Alvarez J, Laskin J and Cooks RG, Phys. Chem. Chem. Phys. **7**, 1490, 2005.
10. Gologan B, Wiseman JM and Cooks RG, Ion Soft Landing: Instrumentation, Phenomena, and Applications. In Principles of Mass Spectrometry Applied to Biomolecules; Laskin J and Lifshitz C, Eds.; John Wiley & Sons., Inc.: Hoboken, NJ, 2006.
11. Hanley L and Sinnott SB, Surf. Sci. **500**, 500, 2002.
12. Tseng AA, Small **1**, 594, 2005.
13. Jacobs DC, Annu. Rev. Phys. Chem. **53**, 379, 2002.
14. Cooks RG, Terwilliger DT, Ast T, Beynon JH and Keough T, J. Am. Chem. Soc. **97**, 1583, 1975.
15. Cooks RG, Ast T and Mabud A, Int. J. Mass Spectrom. Ion Processes **100**, 209, 1990.
16. Williams ER, Fang LL and Zare RN, Int. J. Mass Spectrom. Ion Processes **123**, 233, 1993.
17. Morris JR, Kim G, Barstis TLO, Mitra R and Jacobs DC, J. Chem. Phys. **107**, 6448, 1997.
18. Gu CG, Somogyi A, Wysocki VH and Medzihradszky KF, Anal. Chim. Acta **397**, 247, 1999.
19. Pradeep T, Shen JW, Evans C, and Cooks RG, Anal. Chem. **71**, 3311, 1999.
20. Pradeep T, Evans C, Shen JW, Cooks RG, J. Phys. Chem. B **103**, 5304, 1999.
21. Smith DL, Selvan R and Wysocki VH, Langmuir **19**, 7302, 2003.

22. Vincenti M and Cooks RG, Org. Mass Spectrom. **23**, 317, 1988.
23. Park SC, Maeng KW, Pradeep T and Kang H, Angew. Chem. Int. Ed., **40**, 1497, 2001.
24. Douglas DJ, Org. Mass Spectrom. **17**, 198, 1982.
25. Bier ME, Vincenti M and Cooks RG, Rapid Commun. Mass Spectrom. **1**, 92, 1987.
26. Cooks RG, Ast T, Pradeep T and Wysocki V, Acc. Chem. Res. **27**, 316, 1994.
27. Franchetti V, Solka BH, Baitinger WE, Amy JW and Cooks RG, Int. J. Mass Spectrom. Ion Processes **23**, 29, 1977.
28. Miller SA, Luo H, Pachuta SJ and Cooks RG, Science **275**, 1447, 1997.
29. Mayer PS, Tureček F, Lee HN, Scheidemann AA, lney TN, Schumacher F, Štrop P, Smrčina M, Pátek M and Schirlin D, Anal. Chem. **77**, 4378, 2005.
30. Volný M and Tureček F, J. Mass Spectrom. **41**, 124, 2006.
31. Siuzdak G, Hollenbeck T and Bothner B, J. Mass Spectrom. **34**, 1087, 1999.
32. Claeyssens F, Pratontep S, Xirouchaki C and Palmer RE, Nanotechnology **17**, 805, 2006.
33. Kaiser B, Bernhardt TM, Stegemann B, Opitz J and Rademann K, Phys. Rev. Lett. **83**, 2918, 1999.
34. Messerli S, Schintke S, Morgenstern K, Sanchez A, Heiz U and Schneider WD, Surf. Sci. **465**, 331, 2000.
35. Mitsui M, Nagaoka S, Matsumoto T and Nakajima A, J. Phys. Chem. B **110**, 2968, 2006.
36. Neuendorf R, Palmer RE and Smith R, Chem. Phys. Lett. **333**, 304, 2001.
37. Palmer RE, Pratontep S and Boyen HG, Nat. Mater. **2**, 443, 2003.
38. Yamaguchi W, Yoshimura K, Tai Y, Maruyama Y, Igarashi K, Tanemura S and Murakami J, Chem. Phys. Lett. **311**, 341, 1999.
39. Alvarez J, Cooks RG, Barlow SE, Gaspar DJ, Futrell JH and Laskin J, Anal. Chem. **77**, 3452, 2005.
40. Alvarez J, Futrell JH and Laskin J, J. Phys. Chem. A **110**, 1678, 2006.
41. Gologan B, Takats Z, Alvarez J, Wiseman JM, Talaty N, Ouyang Z and Cooks RG, J. Am. Soc. Mass. Spectrom. **15**, 1874, 2004.
42. Ouyang Z, Takats Z, Blake TA, Gologan B, Guymon AJ, Wiseman JM, Oliver JC, Davisson VJ and Cooks RG, Science **301**, 1351, 2003.
43. Volný M, Elam WT, Branca A, Ratner BD and Tureček F, Anal. Chem. **77**, 4890, 2005.
44. Gao L, Lyn ME, Bergeron DE and Castleman AW, Int. J. Mass Spectrom. **229**, 11, 2003.
45. Loffler D, Jester SS, Weis P, Bottcher A and Kappes MM, J. Chem. Phys. **124**, 054705, 2006.
46. Nanita SC, Takats Z and Cooks RG, J. Am. Soc. Mass. Spectrom. **15**, 1360, 2004.
47. Rader HJ, Rouhanipour A, Talarico AM, Palermo V, Samori P and Mullen K, Nat. Mater. **5**, 276, 2006.
48. Cowin JP, Tsekouras AA, Iedema MJ, Wu K and Ellison GB, Nature **398**, 405, 1999.
49. Tsekouras AA, Iedema MJ and Cowin JP, J. Chem. Phys. **111**, 2222, 1999.
50. Arenz M, Landman U and Heiz U, Chemphyschem **7**, 1871, 2006.
51. Nagaoka S, Matsumoto T, Okada E, Mitsui M and Nakajima A, J. Phys. Chem. B **110**, 16008, 2006.
52. Nagaoka S, Matsumoto T, Ikemoto K, Mitsui M and Nakajima A, J. Am. Chem. Soc. **129**, 1528, 2007.
53. Feng BB, Wunschel DS, Masselon CD, Pasa-Tolic L and Smith RD, J. Am. Chem. Soc. **121**, 8961, 1999.
54. Siuzdak G, Bothner B, Yeager M, Brugidou C, Fauquet CM, Hoey K and Chang CM, Chem. Biol. **3**, 45, 1996.
55. Miller SA, Luo H, Jiang X, Rohrs HW and Cooks RG, Int. J. Mass Spectrom. Ion Processes **160**, 83, 1997.
56. Pradeep T, Feng B, Ast T, Patrick JS, Cooks RG and Pachuta SJ, J. Am. Soc. Mass. Spectrom. **6**, 187, 1995.
57. Shen JW, Grill V, Evans C and Cooks RG, J. Mass Spectrom. **34**, 354, 1999.

58. Wade N, Pradeep T, Shen JW and Cooks RG, Rapid Commun. Mass Spectrom. **13**, 986, 1999.
59. Wade N, Gologan B, Vincze A, Cooks RG, Sullivan DM and Bruening ML, Langmuir **18**, 4799, 2002.
60. Evans C, Wade N, Pepi F, Strossman G, Schuerlein T and Cooks RG, Anal. Chem. **74**, 317, 2002.
61. Shen JW, Evans C, Wade N and Cooks RG, J. Am. Chem. Soc. **121**, 9762, 1999.
62. Volný M, Elam WT, Ratner BD and Tureček F, Anal. Chem. **77**, 4846, 2005.
63. Wang P, Hadjar O, Laskin J, J. Am. Chem. Soc. **129**, 8682, 2007.
64. Nowak P, McIntyre NS, Hunter DH, Bello I and Lau WM, Surf. Interface Anal. **23**, 873, 1995.
65. Ada ET, Kornienko O and Hanley L, J. Phys. Chem. B **102**, 3959, 1998.
66. Park KH, Kim BC and Kang H, J. Chem. Phys. **97**, 2742, 1992.
67. Rabalais JW and Kasi S, Science **239**, 623, 1988.
68. Love JC, Estroff LA, Kriebel JK, Nuzzo RG and Whitesides GM, Chem. Rev. **105**, 1103, 2005.
69. Volný M, Sengupta A, Wilson CB, Swanson BD, Davis EJ and Tureček F, Anal. Chem. **79**, 4543, 2007.
70. Baek DH, Kang H and Chung JW, Phys. Rev. B **49**, 2651, 1994.
71. Bello I, Chang WH and Lau WM, J. Vac. Sci. Tech. **12**, 1425, 1994.
72. Lau WM and Kwok RWM, Int. J. Mass Spectrom. **174**, 245, 1998.
73. Kweon HK and Hakansson K, Anal. Chem. **78**, 1743, 2006.
74. Liang SS, Makamba H, Huang SY and Chen SH, J. Chromatogr. A **1116**, 38, 2006.
75. Wijesundara MBJ, Hanley L, Ni BR and Sinnott SB, Proc. Natl Acad. Sci. USA. **97**, 23, 2000.
76. Wijesundara MBJ, Ji Y, Ni B, Sinnott SB and Hanley L, J. Appl. Phys. **88**, 5004, 2000.
77. Hadjar O, Wang P, Futrell JH, Dessiaterik Y, Zhu Z, Cowin JP, Iedema MJ and Laskin J, Anal. Chem. **79**, 6566, 2007.
78. Vestal ML, Chem. Rev. **101**, 361, 2001.
79. Cooks RG, Ouyang Z, Takats Z and Wiseman JM, Science **311**, 1566, 2006.
80. Fenn JB, Mann M, Meng CK, Wong SF and Whitehouse CM, Mass Spectrom. Rev. **9**, 37, 1990.
81. Raffaelli A and Saba A, Mass Spectrom. Rev. **22**, 318, 2003.
82. Stults JT, Curr. Opin. Struct. Biol. **5**, 691, 1995.
83. Barber M, Bordoli RS, Elliott GJ, Sedgwick RD and Tyler AN, Anal. Chem. **54**, A645, 1982.
84. Bromann K, Felix C, Brune H, Harbich W, Monot R, Buttet J and Kern K, Science **274**, 956, 1996.
85. Geiger RJ, Melnyk MC, Busch KL and Bartlett MG, Int. J. Mass Spectrom. **183**, 415, 1999.
86. Luo H, Miller SA, Cooks R G and Pachuta SJ, Int. J. Mass Spectrom. **174**, 193, 1998.
87. Shen JW, Yim YH, Feng BB, Grill V, Evans C and Cooks RG, Int. J. Mass Spectrom. **183**, 423, 1999.
88. Hadjar O, Futrell JH and Laskin J, J. Phys. Chem. C **111**, 18220, 2007.
89. Laskin J, Wang P, Hadjar O, Futrell JH, Alvarez J and Cooks RG, Int. J. Mass Spectrom. **265**, 237, 2007.
90. Wang H, Chen SF, Li LY and Jiang SY, Langmuir **21**, 2633, 2005.
91. Surrey AR, Name Reactions in Organic Chemistry; Academic Press: New York, 1961.
92. Chan CM, Ko TM and Hiraoka H, Surf. Sci. Rep. **24**, 3, 1996.
93. Gerenser LJ, In Plasma Surface Modification of Polymers; Strobel M, Lyons C, Mittal KL, Eds.; VSP: Zeist, The Netherlands, 1994; 43.
94. Johnston EE and Ratner BD, J. Electron. Spectrosc. Relat. Phenom. **81**, 303, 1996.
95. Yasuda H, Plasma Polymerization; Academic: New York, 1985.
96. Han LM, Timmons RB, Bogdal D and Pielichowski J, Chem. Mater. **10**, 1422, 1998.
97. Hopkins J and Badyal JPS, J. Phys. Chem. **99**, 4261, 1995.
98. Labelle CB and Gleason KK, J. Vac. Sci. Tech. **17**, 445, 1999.

99. Gengenbach TR and Griesser HJ, Surf. Interface Anal. **26**, 498, 1998.
100. Leggett GJ and Vickerman JC, Anal. Chem. **63**, 561, 1991.
101. Wijesundara MBJ, Fuoco E and Hanley L, Langmuir **17**, 5721, 2001.
102. Schaub R, Jodicke H, Brunet F, Monot R, Buttet J and Harbich W, Phys. Rev. Lett. **86**, 3590, 2001.
103. Vajda S, Winans RE, Elam JW, Lee BD, Pellin MJ, Seifert S, Tikhonov GY and Tomczyk NA, Top. Catal. **39**, 161, 2006.
104. Kaiser B, Bernhardt TM, Stegemann B, Opitz J and Rademann K, Nucl. Instrum. Methods Phys. Res., Sect. B **157**, 155, 1999.
105. Schmitz-Hubsch T, Sellam F, Staub R, Torker M, Fritz T, Kubel C, Mullen K and Leo K, Surf. Sci. **445**, 358, 2000.
106. Samori P, Keil M, Friedlein R, Birgerson J, Watson M, Mullen M, Salaneck WR and Rabe JP, J. Phys. Chem. B **105**, 11114, 2001.
107. Bottcher A, Weis P, Jester SS, Loffler D, Bihlmeier A, Klopper W and Kappes MM, Phys. Chem. Chem. Phys. **7**, 2816, 2005.
108. Volný M, Elam WT, Ratner BD and Tureček F, J. Biomed. Mater. Res. B **80B**, 505, 2007.
109. Wang P, Hadjar O, Gassman PL and Laskin J, Phys. Chem. Chem. Phys. **10**, 1512, 2008.
110. Ruoslahti E, Annu. Rev. Cell Dev. Biol. **12**, 697, 1996.
111. Laskin J, Wang P and Hadjar O, Phys. Chem. Chem. Phys. **10**, 1079, 2008.
112. Wade N, Evans C, Jo SC and Cooks RG, J. Mass Spectrom. **37**, 591, 2002.
113. Sullivan TP and Huck WTS, Eur. J. Org. Chem. **2003**, 17, 2003.
114. Corbett AD and Gleason JL, Tetrahedron Lett. **43**, 1369, 2002.
115. Yan L, Marzolin C, Terfort A and Whitesides GM, Langmuir **13**, 6704, 1997.
116. Houseman BT, Gawalt ES and Mrksich M, Langmuir **19**, 1522, 2003.
117. Lee JK, Lee KB, Kim DJ and Choi IS, Langmuir **19**, 8141, 2003.
118. Wagner P, Hegner M, Kernen P, Zaugg F and Semenza G, Biophys. J. **70**, 2052, 1996.
119. Laskin J, Denisov EV, Shukla A, Barlow SE and Futrell JH, Anal. Chem. **74**, 3255, 2002.

Part II
Basic Ion-Matter Interactions in Nanometre Scale Materials

Ion beams have wide applications in the processing and characterisation of materials. These applications all rely on the basic interactions between an energetic ion and the electrons and nuclei of the atoms of the material. Ion beams interacting with bulk material induce changes on a nanometre scale by, for example, localised rearrangement of atoms. The energy that is transferred from the ion to the material can range from many hundreds of keV to milli-electron volts. The outcome of this energy transfer introduces nanometre-scale changes over a wide span of time scales ranging from 10^{-16} s for knock-out of a surface atom to many millenia for the diffusion of actinide elements in oxide matrices that are of importance in repositories for nuclear wastes.

In this part of the book the basic interaction processes are introduced in some detail. This is to provide an underlying fundamental framework for the applications which are discussed in subsequent chapters. This part also contains chapters on modern computational approaches used for theoretical study of ion-matter interactions on a nanometre scale. The idea of this part of the book is to give the reader an introduction to the theoretical approaches and computational methods for prediction and theoretical study for their own applications.

Basics of Ion Scattering in Nanoscale Materials

Harry J. Whitlow and Yanwen Zhang

1 Introduction

Energetic ions interact with materials by collisions with the nuclei and electrons of the atoms that make up the material. In these collisions, energy and momentum is transferred from the projectile particle, which is a moving atom or ion, to the target particles (atomic nucleus or electron). Each collision leads to a slowing down of the moving projectile and also a deflection of the trajectory that gives rise to the term *scattering*, which is often used synonymously to describe the energy transfer process.

In this chapter, we introduce from an experimental viewpoint the underlying theory for ion-solid interactions for analysis and modification of nanometre scale materials. A more detailed theoretical overview of the topic can be found in the recent monographs by Sigmund [1, 2]. Detailed derivations of the formulae introduced will not be given here but can be found in standard texts that are indicated by references.

The treatment here starts by considering an individual scattering event. The results are then used to consider the effects in thick targets where the primary ion encounters a large number of collisions. Subsequently, the primary effects are considered in nanometre materials, which approach the thin-medium limit where the primary particles encounter only a limited number of scattering centres. Finally, the dissipation of the energy deposited by the primary projectiles in secondary processes such as cascades of displaced atoms and electrons will be considered in the thick and thin medium limits. This approach was chosen because it builds upon the standard concepts in ion-solid interactions that are well known and have been widely used in experimental measurements of the stopping force and applications such as Rutherford backscattering spectrometry (RBS), ion beam modification of materials, etc.

H.J. Whitlow (✉)
Department of Physics, University of Jyväskylä, Jyväskylä, FIN-40014, Finland
e-mail: harry.j.whitlow@jyu.fi

2 Binary Scattering

2.1 Energy Transfer in Binary Scattering

The interactions between an energetic ion and target nuclei can be described by collisions between massive point particles. Generally, in ion-solid interactions, it is convenient to treat the scattering interaction between two particles at a time (binary scattering). If energy and momentum are conserved, the collision is termed elastic. On the other hand, the collision is inelastic if there is a conversion between kinetic energy and potential energy. An example of this is where electrons are excited to higher-lying shells during ion-atom collisions.

   Figure 1 schematically illustrates a binary elastic collision. In the collision, energy and momentum are transferred from the projectile particle to the target particle. As the collision is elastic, energy and momentum are conserved:

$$E_1 = E_0 K = E_0 \left\{ \frac{\cos\theta \pm \sqrt{(M_1/M_2)^2 - \sin^2\theta}}{1 + (M_1/M_2)} \right\}^2 \tag{1}$$

$$E_2 = E_0 \Lambda = E_0 \frac{4M_1 M_2}{(M_1 + M_2)^2} cos^2\phi \tag{2}$$

Here K and Λ are the kinematic factors, which represent the fraction of energy retained by the projectile after collision and the energy transferred to the recoiling target particle, respectively. $K + \Lambda = 1$. For $M_1 \leq M_2$ only the positive root is valid. Analytic handling of inelastic and elastic binary scattering is most
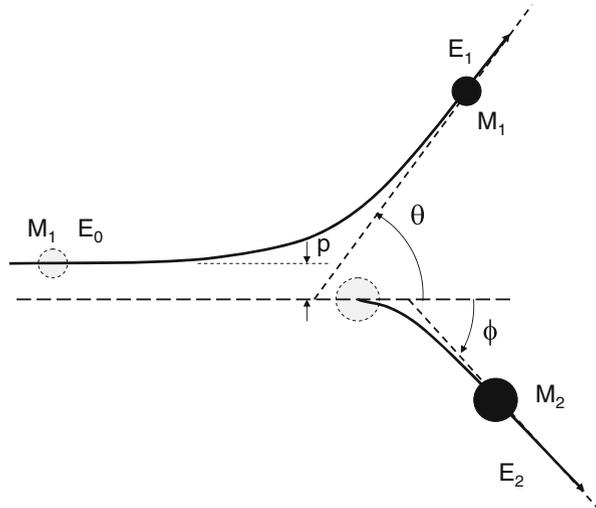

Fig. 1 Binary scattering in the laboratory frame. Following convention, the scattered particle is denoted by subscript 1, while the recoiling target particle is denoted by subscript 2. The initial positions of the particles are denoted by grey toned dashed circles

conveniently done in the barycentric, or centre of mass frame, which is explained in many standard nuclear physics and other texts [1–4].

2.2 Scattering and Scattering Cross-Sections

The *statistical cross-section* is an important concept that is fundamental to ion-matter interactions. The probability of scattering taking place at a single scattering centre is described by σ, the statistical cross-section. For N_i particles incident on a layer of thickness t, containing a volume density of N_0 randomly distributed scattering centres (that may be target nuclei or electrons), the number Y of all scattering events from the centres is:

$$Y = N_i N_0 \sigma t. \tag{3}$$

Generally, we are interested in the probability of a particle being scattered at an angle θ into a solid angle $\delta\Omega$, or the number of scatterings where a certain amount of energy T to $T + \delta T$ is transferred. It is then more useful to use the differential cross-sections for angular scattering and energy transfer, which are, respectively:

$$Y = N_i N_0 t \frac{d\sigma\,(\theta)}{d\Omega} \delta\Omega \tag{4}$$

$$Y = N_i N_0 t \frac{d\sigma\,(T)}{dT} \delta T. \tag{5}$$

The differential cross-section concept can be extended to arbitrary higher differentials; for example, $d^3\sigma\,(\theta, T, \Omega)\,/\,(d\theta dT d\Omega)$ describes the probability of scattering at an angle θ into solid angle $d\Omega$, with energy transfer T.

Another important parameter for the description of scattering is the *impact parameter*, as shown in Fig. 1. If the particle is not deflected in the collision, the impact parameter p would be the distance of closest approach. The impact parameter and the interaction potential $V(r)$ between the point particles determine the scattering and energy loss cross-sections, Eq. (5). Detailed descriptions of the fundamental scattering process are given in a number of texts [1–6].

2.3 Interaction Potentials

The *Coulomb potential* describes the force between two point electrostatic charges $(q_1 e, q_2 e)$ separated by a distance (r). Note that the force is repulsive and isotropic, and the potential is given by:

$$V\,(\mathbf{r}) = \frac{q_1 q_2 e^2}{4\pi \epsilon_0 r}. \tag{6}$$

This is realistic for the force between an ion with charge q_1 and an electron, or between two nuclei with charges Z_1e and Z_2e that are very close so that orbiting electrons do not penetrate the space between them. In this case the scattering probability is described by the differential cross-section, which without too much effort can be obtained analytically [7].

$$\frac{d\sigma}{d\Omega} = \left(\frac{Z_1 Z_2 e^2}{16\pi\epsilon_0} E_0 \right)^2 \left(\frac{4}{sin^4\theta} \right) \frac{\left[(M_2^2 - M_1^2 sin^2\theta)^{\frac{1}{2}} + M_2 cos\theta \right]^2}{(M_2^2 - M_1^2 sin^2\theta)^{\frac{1}{2}}}. \qquad (7)$$

This situation where scattering takes place in a pure Coulomb field is termed *Rutherford scattering*. In general, this condition is well satisfied by backscattered MeV He ions ($< \sim 2.4$ MeV) that form the basis of Rutherford backscattering spectrometry (RBS) . RBS is a widely used near-surface analysis technique.

For lower energy and/or smaller angle collisions, the apsis of the collision is not so close, and electron shells can penetrate into the space between the two nuclei. In this case the negatively charged electrons introduce an electrostatic screening term into equ. 6, with $q_1 = Z_1e$ and $q_2 = Z_2e$.

$$V(\mathbf{r}) = \frac{Z_1 Z_2 e^2}{4\pi\epsilon_0 r} \phi\left(\frac{r}{a} \right). \qquad (8)$$

Here ϕ is the *screening function*; a is termed the *screening length*. The limiting case $\phi(r/a) \overset{r\to 0}{=} 1$ corresponds to Rutherford scattering, while $\phi(r/a) \overset{r\to\infty}{=} 0$ corresponds to the case where the separation approaches infinity so that the nuclear charges Z_1e and Z_2e are completely screened by the electrons.

Various screening lengths, a, have been proposed. For detailed discussions the reader is referred to the standard texts on ion scattering [1, 2, 5, 6].

An important class of potentials are the inverse power-law potentials $V(r) = C_m r^{-1/m}$. A Coulomb potential corresponds to $m = 0.5$ and m becomes progressively smaller ($m = 0.5 \to 0.333 \to 0.25$) as the strength of the screening increases. In these limited range potentials, C_m is a constant. The particular advantage of the power-law potentials is that they are amenable to analytic calculations.

So far, only repulsive-core potentials have been considered. These are applicable when the energies involved greatly exceed the bond energy, so that the latter has had only a negligible effect on the outcome of scattering. For low energy ions, or where moving atoms make collisions with many other atoms simultaneously, a potential that has both attractive- and repulsive-core terms must be used to describe the repulsive and cohesive forces correctly. Depending on the nature of the system these can be isotropic, which is appropriate for close-packed for metal systems, or have terms that describe the bond directionality (e.g., the Tersoff potentials for Si and the Brenner for sp2 and sp3 carbon bonds). These are described in the chapter on Computer Simulation Methods for Defect Configurations and Nanoscale Structures.

3 Energy Loss and Angular Scattering

Consider an ion passing through a thin film whose thickness increases from a single monolayer to a thick layer, but not so thick as to completely stop the ions. When the film is very thin, the ions undergo only one binary collision (*single scattering*). Eventually, as the film thickness increases, some ions are scattered twice (*double scattering*), and then three or more times (*plural scattering*). *Multiple scattering* is often used to describe the situation where the ions are scattered many times. It must be emphasised that for films with thickness on a scale of nanometres, the mean free path between collisions is comparable to or greater than the thickness of the material. Thus, in traversing a nanometre film, only a small fraction of the ions may actually undergo scattering from a nucleus of an atom in the film. This contrasts starkly to the situation often met in conventional ion interaction studies where the media is considered to be sufficiently thick that multiple collisions take place. In dealing with ion interactions in nanometre materials, it is therefore important to consider that we are often dealing with the transition region between a single, or even no scattering and many scattering events. The energy loss and scattering in thick media are often met in many practical situations and therefore this case is considered first. Subsequently, we consider the departures from this situation where single and plural scattering dominate in thin media.

The probability of an ion undergoing n collisions after penetrating a distance x in the target is given by the Poisson distribution [1]:

$$P_n = \frac{(Nx\sigma_A)^n}{n!} \exp(-Nx\sigma_A), \quad n = 1, 2, 3 \ldots \tag{9}$$

where $\sum_{n=0}^{\infty} P_n = 1$. Hence, $P_0 = \exp(-Nx\sigma_A)$, which is the Lambert and Beer law for adsorption of quanta; $P_1 = (Nx\sigma_A)\exp(-Nx\sigma_A)$ corresponds to single scattering and $P_2 = \frac{1}{2}(Nx\sigma_A)^2\exp(-Nx\sigma_A)$ to double scattering, etc. As the projectile passes through the surface, $P_0 = 1$, the possibility for scattering to occur is turned on. It follows that the projectiles penetrate some distance before scattering the mean distance being the mean free path $\lambda = 1/(N\sigma_A)$ [1]. Here σ_A is the cross-section for the particular process in question. It should be noted that the different cross-sections for different processes like ion-electron and ion-nucleus scattering imply that the mean free paths are different for the different processes.

3.1 Penetration of Thick Media

The stopping cross-section, ϵ, is given by $\Delta E = Nx\epsilon$, where Nx is thickness of media in terms of atoms per unit area. Then ϵ represents the energy loss per atom per unit area and is given by:

$$\epsilon = \int T\, d\sigma\,(T). \tag{10}$$

The cross-section $\sigma(T)$ for energy transfer T is determined by the interaction potential discussed above. ϵ, which has units of energy per atom per unit area, is comprised of contributions from ion-electron scattering (*electronic stopping*) and ion-nuclear scattering (*nuclear stopping*).

The stochastic nature of the collisions gives a distribution of ΔE. According to the Poisson distribution [1], $\overline{(n - \langle n_j^2 \rangle)} = \langle n \rangle$, the variance of the energy loss (i.e., the *energy loss straggling*) is:

$$\Omega^2 = N \Delta x \int T^2 d\sigma(T). \tag{11}$$

A related quantity is the net angular deflection, α, of an ion relative to the incident beam direction:

$$\alpha = \int \phi d\sigma(\phi), \tag{12}$$

where ϕ represents the individual deflections of the ions: The average deflection $\langle \alpha \rangle$ will be zero because of the cylindrical symmetry of the collision [1], while the mean square deflection angle is:

$$\langle \alpha^2 \rangle = \overline{(\alpha - \langle \alpha \rangle)^2} = N \Delta x \int \phi^2 d\sigma(\phi). \tag{13}$$

Although contributing little to the energy loss, nuclear scattering does lead to a large number of small-angle scattering events resulting in lateral and angular spreading. After penetrating some distance in to the materials, the angular (α) and lateral spreading (ρ), as indicated in Fig. 2, contribute significantly to the degradation of depth resolution in scattering-based profiling methods, and limit the lateral resolution in nuclear microscopy and MeV ion beam lithography.
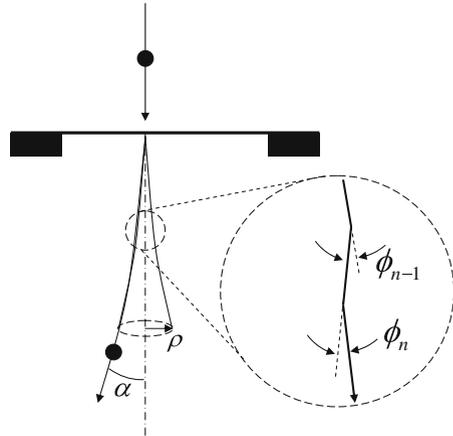

Fig. 2 Angular and lateral spreading distribution of an ion entering matter

When the material is sufficiently thick that a statistical equilibrium is reached, the average deflection is:

$$\langle \alpha^2 \rangle = N \Delta x \int \phi^2 d\sigma. \tag{14}$$

The half-width of the lateral spreading $\rho_{1/2}$ associated with the angular spreading $\alpha_{1/2}$ can be conveniently estimated using the law proposed by Sigmund, Marwick, and Winterbon et al. [8–12]:

$$\rho_{1/2} = x \alpha_{1/2} \Gamma \tag{15}$$

$$\Gamma = 1.77 + 0.172 \tau^{-0.335} \tag{16}$$

with the reduced depth of penetration $\tau = \pi a^2 N x$, and a is the screening length introduced in (Eq. 8).

Matter is constituted of atoms, which from the standpoint of this book can be considered to be made up of negatively charged electrons and positively charged massive point nuclei. If the number of target atoms per unit volume is N_2, the ion can scatter from $N_2 x$ nuclei and $Z_2 N_2 x$ electrons in traversing a distance x in the target medium. If the collisions with nuclei and electrons are separable, we have:

$$\Delta E = N_2 x \epsilon = N_2 x Z_2 \epsilon_{el} + N_2 x \epsilon_n, \tag{17}$$

where ϵ_{el} and ϵ_n are the stopping cross-sections for electronic stopping in terms of single electrons and nuclei, respectively. Provided that the electronic and nuclear stopping processes of the ions in the pure elements are exactly the same as in the compound $A_n B_m$, we have:

$$\Delta E = \frac{m \Delta E_A}{m + n} + \frac{n \Delta E_B}{m + n} = N_2 x \left(\frac{n \epsilon_A + m \epsilon_B}{(m + n)} \right). \tag{18}$$

The right hand relation in Eq. (18) is *Bragg's rule*, which is used for calculating the stopping force for compounds.

The electronic and nuclear stopping forces dominate over different energy regions. Figure 3 illustrates the calculated stopping force for different energy regions. For low energies (few keV), nuclear stopping dominates; while at high energies (MeV), the stopping force is dominated by the contribution from electronic processes. Although the electronic stopping completely dominates above ~ 5 keV, the nuclear stopping contribution has a monotonically decreasing persistent tail. As noted above, the nuclear stopping contribution dominates the angular spreading of MeV beams. This angular spreading has an important influence on ion beam penetration of matter on a nanometre scale. This phenomenon plays an important part in MeV ion beam lithography, ion track writing, and the resolution of RBS and elastic recoil detection analysis, which are discussed in subsequent chapters.

Fig. 3 Electronic and nuclear contributions to the stopping cross-section ϵ, calculated using the SRIM code [5, 6]
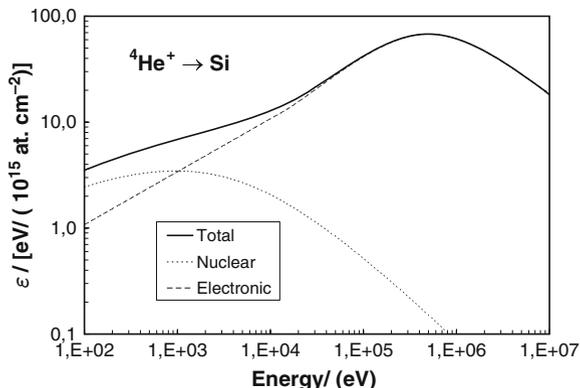
3.1.1 Electronic Stopping

The electronic stopping cross-section in SI units is [2]:

$$\epsilon_e = \frac{Z_1^2 Z_2 e^4}{4\pi \epsilon_0^2 m_e v_1^2} L. \tag{19}$$

In this formulation, all the quantum mechanical behaviour is lumped into the stopping number L. All the other terms are classical, with v_1 the ion velocity. For fast light ions, (e.g., MeV protons and ^{4}He) where 1st-order perturbation theory is valid, the stopping number is the well know, Bethe form L_{Bethe}; while for slow heavy ions the classical Bohr value L_{Bohr} is more appropriate [1].

$$L_{Bethe} = log_e \left(\frac{2m_e v_1^2}{I} \right), \quad L_{Bohr} = log_e \left(\frac{1.12292 e_e v_1^3}{Z_1 e^2 \omega} \right). \tag{20}$$

Here $I = \hbar\omega$ is the ionisation potential. For practical purposes the ZBL (Ziegler, Biersack, Littmark [1]) electronic stopping cross-section is used. This is a variation of the Brandt-Kattawa theory. In ZBL, Bethe's stopping number L_{Bethe} is assumed as the basis of a semi-empirical approach where the data for a wide range of ions and targets is fitted to an effective charge γ by scaling to the stopping cross-section for protons ϵ_H. Then the stopping cross-section scales for ions heavier than H according to the effective charge:

$$\epsilon_{HI} = \epsilon_H Z_1^2 \gamma^2. \tag{21}$$

A good starting point for the effective charge fraction is the Northcliffe value:

$$\gamma^2 = 1 - exp \left[\frac{-v_1}{\left(v_0 Z_1^{2/3} \right)} \right]. \tag{22}$$

ZBL use a more sophisticated approximation that takes into account the velocity relative to the Fermi velocity of the target electrons. In considering penetration of ions through nanometre material, it should be borne in mind that the ion may enter with a nonequilibrium charge state and gradually acquire a charge state that approaches the equilibrium by electron capture and loss in individual collisions [13]. The implication of the finite mean free path is that in nanometre materials, the stopping may actually differ from the equilibrium bulk values because the charge state is approaching, but has not actually reached, the equilibrium charge state condition.

3.1.2 Nuclear Stopping

Nuclear stopping involves scattering of the ions by the nuclei of the target atoms. In this process, energy is transferred to the target nuclei that recoil, pulling their electrons with them. The energy transferred can be sufficient to move the target out of its normal position by breaking bonds, and in doing so lead to so-called *displacement damage*. The nuclear stopping dominates at low energy (e.g., Fig. 3); however, it is persistent even at high (MeV) energies where the dominating mode of energy loss is electronic stopping. Using a power law potential, the nuclear stopping cross-sections can be expressed in the form:

$$\epsilon_n = \frac{\lambda}{2\,(m-1)}\tilde{\epsilon}^{\,1-2m}. \tag{23}$$

Here m defines the inverse power (discussed above), $\tilde{\epsilon}$ is a reduced energy and λ is a constant [1].

3.1.3 Ion Ranges

Energetic ions will, penetrating matter, deposit energy in the material by scattering. The depth they penetrate is called the range R:

$$R = \int_{E_0}^{0} dx\,\frac{1}{Nx\epsilon\,(E)}. \tag{24}$$

This assumes that the stopping takes place continuously, so a statistical quasi-equilibrium is reached. The stochastic nature of the scattering leads to the range being a distributed quantity with a variance ΔR that is termed the *range straggling*. The range distributions are discussed in Box 3.

3.2 Energy Loss and Scattering in Nanometre-Scale Materials

An important difference between material on a nm and μm scale is that the number of atoms encountered by ions becomes small. This can be understood by considering that 1 nm of Si corresponds to 3.7 monolayers, while 1 μm corresponds to 3700 monolayers. Ions with energies $>$ 10 eV essentially only interact along their path

with a single atom at a time; hence the number of scattering events that can take place in penetrating on a scale of nm is extremely limited.

The surface may be considered as a singularity. Above the surface the ion does not undergo any scattering, while as it traverses the surface the scattering processes are abruptly turned on. Deep in the media a statistical quasi-equilibrium is set up and the individual ions approach the statistical average behaviour. The intermediate layer where the quasi-equilibrium is approached represents a transition region. The transition region thickness depends through Eq. (9) on the cross-section.

3.2.1 Energy Loss in Thin Media

In the case considered above, where the ion penetrates a target of moderate thickness, it undergoes a large number of collisions. Under these conditions, the energy transfer $T_{max} << \Omega$, which leads to a Gaussian energy loss profile. A small increase in target thickness then leads to an energy loss distribution that is characterised by a shifted Gaussian. Under these conditions, the central limit theorem is valid [14]. If, on the other hand, the target is so thin that the maximum energy transfer in a single scattering event is large compared to the distribution width ($T_{max} >> \Omega$), a small increase in thickness will lead to a change in both the mean energy loss and, more significantly, the shape of the energy loss distribution, since the condition for the central limit theorem to apply is no longer satisfied. Under these conditions the energy loss is a multiple convolution of the energy loss probabilities in a single collision. The number of convolutions involved corresponds to the number of scattering events. The energy loss distribution is governed by the Bothe-Landau formula, which describes the statistics of cumulative scattering events. A discussion of this is theoretically challenging [2] and outside the scope of this book. However, we will highlight the importance of these phenomena by some important cases.

One important limit is that where the target is extremely thin, the energy transferred is much smaller than the average energy loss. In this case, for a cross-section that is described by a pure-Coulomb interaction, the distribution will tend to a Landau distribution [1]. This is characterised by a narrow peak with a persistent tail, originating from the small number of collisions where a large amount of energy is transferred. As the stopping media thickness increases, the energy loss distribution will smoothly transform from the Landau case to the Gaussian thick media case. In the intermediate case, the energy loss distribution will be described by the Vavilov distribution [1, 30]. The practical implication for high-resolution depth profiling in nanometre materials is that the energy loss distribution may not be simply considered to be a Gaussian energy loss distribution. Under these conditions, the mean and modal energy loss will deviate at small thicknesses but converge as the thickness of the stopping media increases. The Vavilov and Landau distributions are nowhere near as straightforward to use and interpret as a Gaussian because exact evaluation requires complex integration. A good way forward in these cases seems to be fitting the experimental spectra with energy loss distributions using Monte Carlo calculations, taking into account the energy loss processes for each scattering event. An

alternative approach would be to use numerical convolution of the single collision energy loss spectrum.

3.2.2 The Approach to Equilibrium Charge States

Ions produced from accelerators generally have fixed charge states that are defined by electrostatic and magnetic optical elements. As the ions traverse the surface, electron transfer between the ion and the target media starts to take place. This capture and loss of electrons by the ions leads to a charge state that is dynamically changing in discrete steps, but tends towards a net equilibrium charge state, which depends on the ion atomic number Z_1 and the velocity of the ion, according to Eq. (22). Thus, as the stopping depends on the charge state of the ion in the media, until the quasi-equilibrium is set up, the stopping force encountered by the ion in the surface region will be enhanced or suppressed relative to the equilibrium state.

3.2.3 Ion Implantation Doping on Nanometre Materials

In microelectronic technology, doping of semiconductors by ion implantation has long been a mainstay of manufacturing. As the sizes of devices shrink now to nanometre dimensions, the number of doping atoms in each part of the device decreases. Typical semiconductor dopant concentrations are 10^{18} ions cm^{-3}, which, for a volume of 1000 nm^3 (about the size of a quantum dot), corresponds to only a single dopant atom. This represents a limit to the use of ion implantation doping because at these concentrations solid solubility limit. This gives the stochastic nature of ion impingement, implies that ion implantation is not easily used for doping of quantum dot structures. An interesting approach pioneered by Jamiesson et al. for fabricating Q-bits is to circumvent these stochastic ion impingement effects by using the substrate as a detector and irradiating till an impinging dopant ion is detected [15].

4 Energy Transfer by Secondary Particle Cascades

4.1 Introduction

Materials subjected to ion bombardment are modified not only by the primary ion being incorporated by ion implantation, but also by the kinetic energy deposited by stopping of the primary ion and recoils leading to a whole plethora of secondary effects. As discussed above, ions interact by electron and nuclear scattering. Figure 4 schematically illustrates the interrelation of these effects, which stem from ion-electron and ion-nucleus scattering. The energy deposited by the ion is either carried out of the material by recoil and sputtered atoms, X-rays, secondary electrons, photons, etc., or eventually ends up in some form of stored energy (e.g., structural defects) and heat.
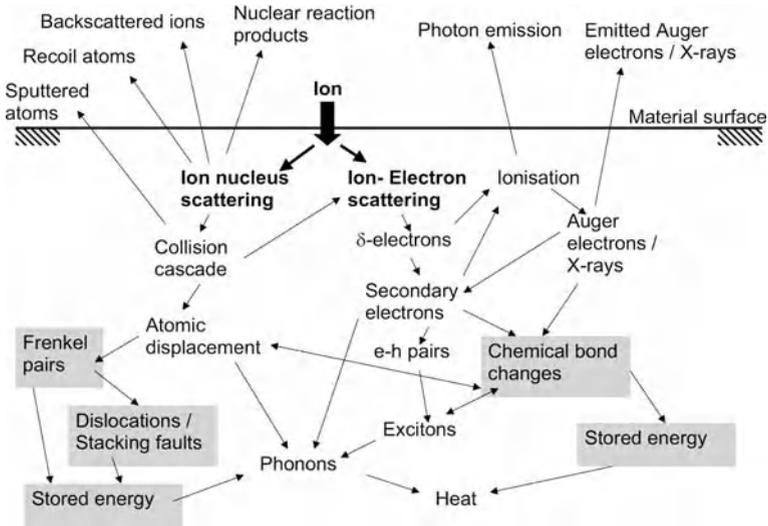
Fig. 4 Schematic illustration of the evolution of the processes taking place in a solid material irradiated with ions. The *grey shaded* processes correspond to long term effects. After Zhang and Whitlow [16]

Although, ions can transfer energy during scattering by classical and quantum mechanical processes, it is convenient to consider the scattering as elastic, or, in the case of bound electrons, quasi-elastic, scattering of free particles at rest (Fig. 1). The fraction of energy transferred in a classical collision given by Eq. (2) can be substantial. This energy is then subsequently deposited in subsequent scattering, as illustrated schematically in Fig. 5, that can extend over many generations to form a *collision cascade*.
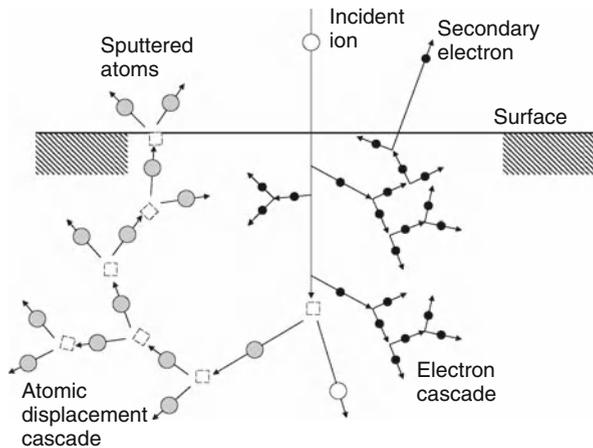
Fig. 5 Schematic illustration of atomic displacement and electron collision cascades brought about by an incident ion penetrating material. The *white circles* denote the incident ions, *grey circles* recoiling target atoms and *dashed squares* vacancies, resulting from ion-nucleus scattering. The *small black circles* denote electrons in the electron cascades

Table 1 Characteristic times for ion induced changes

Process	Timescale
Ion-nucleus scattering	5–500 as
Ion-electron scattering	0.5–100 as
X-ray, Auger electron production	0.1–0.5 fs
Period of lattice thermal vibration	10 fs
Electron cascade	10 fs
Atomic displacement cascade	200 fs–2 ps
Defect diffusion	0.1–2 ns
Thermal heating	100 ns–10 μs
Electron-hole plasma	1–200 ns
Defect clusters and dislocations	ns – stable
Chemical effects	fs – stable

The processes in Fig. 4 take place over different timescales (Table 1). Higher energy processes, such as the ejection of Auger electrons or X-rays, occur over short time scales (fs-ps); while processes that involve energies close to thermal, such as vacancy mediated migration or the cross-linking of a polymer to form a latent image, take place over a longer time scale.

4.2 Atomic Displacement Cascades

In an atomic displacement cascade (Fig. 5), energy is transferred to target nuclei in a binary collision (Fig. 1). Depending on the amount of energy transferred, the recoil energy can be dissipated as heat (phonon production) or, if it exceeds the threshold for displacement, a primary knock-on atom (PKA) can be created, leaving behind a vacancy. Eventually, the PKA loses its kinetic energy and comes to rest in the interstitial position between normally occupied atomic sites. The vacancy and interstitial are termed a *Frenkel pair*. For this reason, the collision cascade is often referred to as an *atomic displacement cascade*. The number of PKAs in a collision cascade can vary widely from a 1 to 2 when the energy transferred is close to the displacement energy to $\sim 10^4$ for 500 keV Pb^+ ion incident on Si. For Rutherford scattering (realistic for MeV ions), the PKA distributions have some characteristic features. The energy distribution $N_{PKA}(E)$ follows closely a $1/E^2$ dependence, while most recoils have direction vectors close to perpendicular.

Figure 6 shows the energy and angular PKA distributions 1 MeV H^+ irradiation. In this case the potential is a screened Coulomb potential (Eq. 8) with the ZBL [6] screening function. The long-range nature of the Coulomb (Eq. 6) and screened Coulomb (Eq. 8) potentials implies that the majority of the PKA are directed near perpendicular to the incident ion direction (Fig. 6a), and from Eq. (2) have low energies. The energy distribution (Fig. 6b) has an energy dependence close to $1/E^2$ for high energies, which is characteristic of Rutherford scattering.

The PKAs slow down in a similar manner as the incident ions by a succession of ion-electron and ion-nucleus scattering. In the continuous slowing down
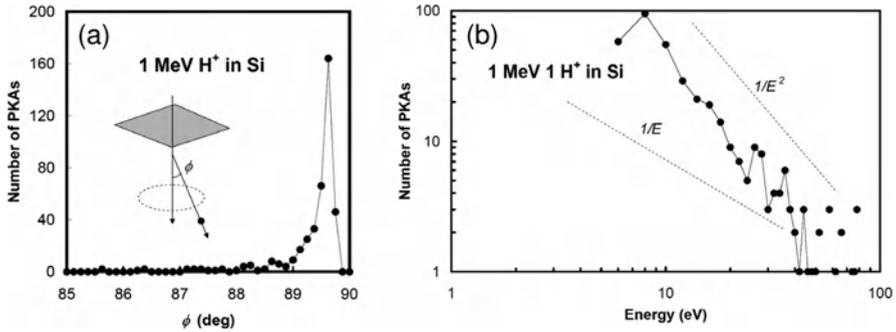
Fig. 6 Angle (**a**) and energy (**b**) distributions of PKA in amorphous Si for 1 MeV H$^+$. After Whitlow and Nakagawa [17]

approximation (CSDA), the successive collisions are considered to be acted upon by mean electronic and nuclear stopping forces. On a nanometre scale; the distances between collisions becomes significant [18] for both electrons and ions. This leads to non-Gaussian energy loss distributions [1] and spatially separated subcascades for both atomic displacements and secondary electrons.

Collision cascades have a number of useful features for nanoenegineering. Firstly, they constitute a "nanoblender" that can mix layers of atoms by *ion beam mixing*. Mixing of different atoms takes place by non-thermal ballistic processes (e.g. the collision cascade mixing which is nonthermal, and the knock-on of atoms in the forward direction). In addition, *radiation enhanced diffusion* may be significant because of the extra mobile defects (interstitials, vacancies) introduced by the ion irradiation. Secondly, as illustrated in Fig. 5, some knock-on atoms may be created sufficiently close to the surface with sufficient energy to overcome the surface binding energy and escape as *sputtered atoms*. Detailed discussion of sputtering is given in Box 2. The maximum sputter escape depth is governed by the knock-on energy and is typically 1–3 atomic layers. The *sputtering* process underlies secondary ion mass spectrometry profiling, focused ion beam technology (chapter "Focused Ion Beam Machining and Deposition") and plasma deposition and etching (chapter "Plasma Etching and Integration with Nanoprocessing"). The sputtering phenomena are rich and wide ranging, and have been reviewed in depth in a series of volumes [19–22].

4.3 Electron Cascades

The electronic stopping of primary ions and energetic secondary recoiling atoms excites core and valence electrons in the target to bound and free states by resonant and nonresonant excitation. For quasi-classical scattering, the energy of free electrons $E_f = T - E_b$, where E_b is the binding energy and $T \approx 4\,(m_e/M_1)\,cos^2\phi$ from Eq. (2). The energy of energetic primary electrons, often termed *δ-electrons*, is transferred in subsequent collisions with electrons in the target material. For primary

ions with MeV and keV energies, the vast majority of secondary electrons deposit their energy giving a dose distribution that falls off very rapidly within 10 nm of the ion track. This is illustrated in Fig. 7a, which shows the radial dose distribution for 1 MeV protons in PMMA (polymethylmethacrylate). The deposited dose at low fluences bond breaking leads to a reduction of the average molecular weight by breaking backbone and side chain bonds. Furthermore, at high doses cross-linking between backbone chains can take place. This affects the ability of solvents to dissolve the polymer, giving it useful properties as a resist material for lithography, such as illustrated in Fig. 7b and discussed in detail in chapter "Proton Beam Writing: A New 3D Nano Lithographic Technique". The small angular deflection of the incident ions and the short transverse spreading of the dose around the ion trajectories combine to give the extremely vertical sidewalls, as shown in Fig. 7b for a pattern written in 8 μm thick PMMA. A concise overview of the effects of irradiation on polymeric materials is given in [23]. The energy deposited by the electrons can induce transient effects like flourescence, as well as having more lasting effects such as the radiolysis which is important in DNA damage in biological materials (chapter "Nanoscale Engineering in Biosciences"). A more detailed description of the nanometre scale electronic effects induced by ions is given elsewhere by the authors [16].
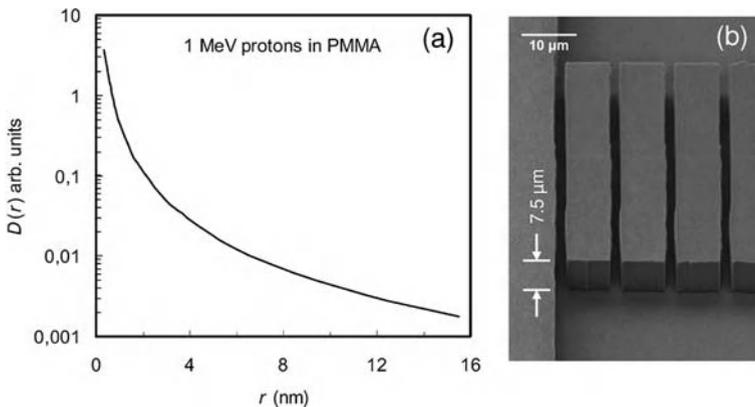

Fig. 7 (**a**) Radial distribution of the dose distribution $D(r)$ around the ion trajectory for a 1 MeV protons in PMMA. After Zhang and Whitlow [16]. (**b**) A rectangular channel with straight vertical edges produced in PMMA using 3 MeV He^{2+} ions. (H.J. Whitlow, N. Puttaraksa, S. Gorelick and T. Sajavara, unpubished data)

4.4 Secondary Cascades in Nanometre Scale Materials

Unlike the case of materials in a μm size scale, the distances between collisions with ions and electrons is comparable in nanometre-scale materials with the thickness of the thin film or other structures. In this connection, the close proximity of the surface can considerably modify the behaviour, compared to "bulk" material far from the

surface. Indeed, the truncation of the collision cascade by the surface allows energetic recoil atoms to escape the surface by sputtering (Fig. 5). From a collisional viewpoint, as the ion penetrates the surface, the probability of undergoing scattering with the nuclei of the target atom increases. The range of the interaction potential determines the point of approach where the scattering probability starts to rise above zero. This point where the interaction starts may actually lie above the surface (e.g., when hollow atoms are created from an impinging highly-charged ions, such as Ar^{12+}, by electron transfer from the surface to high-lying bound states). The influence of the surface is important in a number of practical situations (1) Scattering from apertures where the finite emittance width of practical ion beam results in a flux impinging on the aperture edge at a glancing angle. This gives rise to both an increase in the smearing of the dose profile across the edges of aperture masks [24], as well as a change in the dose at the mask edge [25, 26] in ion implantation. (2) Sputtering where the penetration of the ion to some depth below the surface before creating any PKA results in a net backwards directed recoil flux at the surface, resulting in a nonisotropic sputtered flux [27–29].

5 Future Prospects

It is evident from the above that energetic ions, because of their strong interaction through Coulomb forces, provide the possibility of quantitatively characterising and controllably modifying nanometre scale materials. The significant difference with "bulk" (μm) scale materials is that the mean free paths for ion-electron and ion-nuclei collisions become comparable with the size of the media, and the details of the mesoscopic Poisson statistical behaviour become significant. Rather than assuming that the central limit theorem holds, giving a Gaussian energy loss, a Bothe-Landau approach [1] is needed. Although at first sight, it appears to complicate the situation, the Poisson statistical behaviour opens new possibilities for analysis and modification. This gives the short wavelength (fm-pm) of energetic ions, and trajectories that are extremely straight on an atomic scale. This and the great numbers of electrons along the straight ion track, implies that ions with energies from eV to MeV offer many possibilities for characterisation and modification of nanometre materials. This has already been exploited (chapters "Low and Medium Energy Ion Scattering for Near Surface Structure and Nanoscale Depth Profiling", "Thin Film Characterisation by means of MeV Ion Beams", "Nanoscal Materials Defect Characterisation", and "Focused Ion Beam Machining and Deposition") but could also be exploited further for extremely high resolution imaging. This has already been used for material modification by MeV ion beam lithography (chapter "Proton Beam Writing: A New 3D Nano Lithographic Technique"), and ion track processes (chapter Micro- and Nanoengineering with Ion Tracks), and has potential as the basis of novel analytical and imaging techniques.

From the analysis point of view, the small energy loss in traversing a few nanometres or so of materials implies that extremely high resolution detectors must be used.

Recent advances in mK cooling and SQID (superconducting quantum interference device) technologies open up the possibility of extremely high resolution super-conducting transition detectors that have resolutions on a scale of eV, rather than 10 keV for conventional Si junction detectors. Although these resolutions can be reached with magnetic and electrostatic analysers, calorimetric and time of flight (ToF) spectrometers are charge-state blind, which can overcome difficulties with parameterising charge state changes.

The strong Coulomb interaction also allows for collective interactions with the ions in a crystal. This is the basis of the channelling technique (chapter "Nanoscale Materials Defect Characterisation"), as well as blocking used for location studies in medium energy ion scattering spectroscopy (chapter "Nanomaterials Science with Radioactive Ion Beams"). The channelling technique has been widely used as a crystal defect diagnostic tool, but the ability to steer and collimate ions has as yet not been very much exploited for materials modification. Given the ease with which channelling can be destroyed, it may find application for spatially defined defect engineering.

References

1. P. Sigmund, *Particle penetration and radiation effects*, (Springer, Berlin, 2006)
2. P. Sigmund, *Stopping of heavy ions a theoretical approach*, (Springer, Berlin, 2004)
3. M.T. Robinson, *Sputtering by particle bombardment I* (ed.) R. Behrisch (Springer, Berlin, 1981)
4. M.W. Thompson, *Defects and radiation damage in metals*, (Cambridge Univ. Press, Cmabridge,1969) p. 105–109.
5. J.F. Ziegler et al., SRIM code v2003.26, www.srim.org
6. J.F. Ziegler, J.P. Biersack and U. Littmark, *The stopping and range of ions in solids*, (Pergamon Press, New York, 1985).
7. W.-K. Chu, J.W. Mayer and M.A. Nicolet, *Backscattering Spectrometry*, (Academic Press)
8. P. Sigmund and K.B. Winterbon, Nucl. Instrum. Methods 119 (1974) 541.
9. A.D. Marwick and P. Sigmund, Nucl. Instrum. Methods 126 (1974) 317.
10. G. Amsel, G. Battistig, A. L'Hoir, Nucl. Instrum. Methods B201 (2003) 325.
11. G. Amsel, G. Battistig, A. L'Hoir, Nucl. Instrum. Methods B 219 (2004) 1037.
12. J.R. Bird and J.S. Williams, *Ion beam analysis*, (Academic Press, 1989) p. 620.
13. H.J. Whitlow and H. Timmers, *Electrostatic accelerators*, (ed.) R. Hellborg, (Springer, Berlin, 2005).
14. N. Bohr, Kungl. Dan. Vidensk. Selsk. Mat.-Fys. Medd. 18 (1948) 8.
15. D.N. Jamieson, V. Chan, F.E. Hudson, S.E. Andresen, C. Yang, T. Hopf, S.M. Hearne, C.I. Pakes, S. Prawer, E. Gauja, A.S. Dzurak and R.G. Clark, Nucl. Instrum. Methods B 249 (2006) 221–225.
16. Y. Zhang and H.J. Whitlow, Modification of materials by MeV ion beams. In *Electrostatic accelerators*, (ed) R. Hellborg, (Springer, Berlin Heidelberg New York 2005) pp. 506–529.
17. H.J. Whitlow and S.T. Nakagawa, Nucl. Instrum. Methods B 260 (2007) 468.

18. C.N.B. Udalagama, A.A. Bettiol and F. Watt, Nucl. Instrum. Methods B 260 (2007) 384–389.
19. *Sputtering by particle bombardment I*, (ed.) R. Behrisch, (Springer, Berlin, 1981).
20. *Sputtering by particle bombardment II*, (ed.) R. Behrisch, (Springer, Berlin, 1983).
21. *Sputtering by particle bombardment III*, (ed.) R. Behrisch, (Springer, Berlin, 1991).
22. *Sputtering by particle bombardment: Experiments and computer calculations from threshold to MeV energies*, (ed.) R. Behrisch and W. Ekstein, (Springer, Berlin, 2007).
23. A. Chapiro, Nucl. Instrum. Methods B32 (1988) 111–114.
24. S. Gorelick, T. Sajavaara, M. Laitinen, N. Puttaraksa and H.J. Whitlow, Mater. Res. Soc. Symp. Proc. 1020 (2007) 1020-GC03-04.
25. O.V. Hul'ko, D. Danailov, D. Karpuzov, M. Zinke-Allmang, Phys. B 226 (1999) 345–355.
26. R.C. Wijburg, G.J. Hemink and J. Middelhoek, strongly asymmetric doping profiles at mask edges in high-energy ion implantation. IEEE Trans. Elect. Dev. 37 (1990) 79–87.
27. M.W. Sckerl, P. Sigmund and M. Vicanek, Particle fluxes in atomic collision cascades, Mat. Fys. Medd. 44 (1996) 3
28. H.H. Andersen, B. Stenum, T. Sørensen and H.J. Whitlow, Nucl. Instrum. Methods B6 (1985) 459.
29. M. Hautala and H.J. Whitlow, Nucl. Instrum. Methods B6 (1985) 466.
30. H. Bishel and R.P. Saxon, Phys. Rev. A11 (1975) 1286.

Box 1: Stopping of Ions in Nanomaterials

Yanwen Zhang and William J. Weber

The stopping of ions in solids is due to the energy loss as a result of the resistance to ion passage by the electrons and atomic nuclei in the material. When an ion penetrates a solid, it experiences a number of collisions. Energetic charged particles interact with both electrons and atoms in materials. Kinetic energy transfers to atoms can result in displacement of atoms from their original sites, thereby forming atomic-scale defects in the structure. Energy transfers to the target electrons (either bound or free) produce electron-hole pairs that can result in charging of pre-existing defects, localized electronic excitations, rupture of covalent and ionic bonds, enhanced defect and atomic diffusion, increased free energy, changes in phase transformation dynamics, as well as formation of atomic-scale defects. Such atomic collisions and ionization processes can modify the physical and chemical behavior of nanomaterials. This box discusses irradiation-induced defects, and addresses nanostructure engineering and radiation effects in nanomaterials, as well as the scientific challenges of ion-solid interactions.

1 Irradiation-Induced Defect Dynamics

Since the energy transferred from an ion to an atomic nucleus in a single collision is often many times greater than the binding energy, the atom can be displaced from its site in the structure. Moreover, atoms that are displaced by incident ions may receive so much energy that they in turn can displace other atoms, creating a cascade of atomic collision events. Eventually, a large number of atoms within the solid can be set in motion. This can lead to a distribution of vacancies, interstitial atoms, mixing of layers in the vicinity of interfaces, or lattice disorder in the region around the ion track, etc. In crystalline materials, as the number of incident ions increases, the individual disordered regions can begin to overlap. At some point, a heavily damaged layer is formed. The total amount and type of disorder, as well as

Y. Zhang (✉)
Pacific Northwest National Laboratory, Richland, WA 99352, USA
e-mail: Yanwen.Zhang@pnl.gov

its distribution in depth, depend on ion species, energy, dose, target material, and the irradiation temperature.

Nanostructured materials with building blocks in the 1–100 nm size are a new class of materials that has been making a revolution in the last decade. The properties of materials can be engineered as systems of confined low dimensionality (nanometer size particles of metal, ceramic, or polymers). The ability to alter the physical, electronic, and optical properties by merely changing the size of the nanoparticles is extremely suitable for a variety of applications. Due to their exotic microstructure, exceptional size-dependent properties, and technological implications, nanostructured materials have attracted wide interest for novel catalytic, sensor, membrane, structural, electric, magnetic, and biomaterial applications.

Irradiation-induced defects in nanomaterials are different from bulk materials because of the limited size in one or more directions. Contrary to bulk systems, where all the energy is eventually absorbed, energetic particles may traverse these nanoscaled objects with little energy deposited. Correspondingly, the amount of damage in individual nanostructures is low. On the other hand, the finite size of the nanostructures may affect the electronic structure and thus the mechanisms of conversion of electronic excitations to kinetic energy of the atoms [1].

2 Nanostructure Engineering by Energetic Ion Beams

Irradiating solids with energetic particles is usually thought to introduce disorder, normally an undesirable phenomenon. But recent ion irradiation studies have demonstrated that there are a variety of possibilities in which energetic ions can play a crucial role in the engineering of nanomaterials. Beneficial effects of irradiation on engineering of nanostructure carbon materials have been demonstrated, as in the case of irradiation-mediated engineering, self-assembly, or self-organization in carbon nanosystems. The atomic structure and morphology of carbon nanotubes or related structures can be tailored by irradiation [1, 2], and they can be interconnected or merged in a controllable way [3]. Irradiation can give rise to extreme pressure inside nanotubes [4] or fullerene-like onions [2], so that these systems can be used as compression cells to induce high-pressure transformations of materials on a nanometer scale. Moreover, irradiation can be used to tailor the mechanical, electronic, and magnetic properties of nanostructured carbon. Intensive work on irradiation-assisted nanostructure engineering has also been carried out in noncarbon materials. Nanoscale manipulation of pyrochlore has been demonstrated for the development of mixed ionic- and electronic-conducting ceramic oxide membranes for oxygen separation in fuel cells [5], formation of self-assembled surface ripples, and uniformly distributed metallic nanodots [6]. Recently, ion-induced ripple structures have been of particular interest for the fabrication of nanoscale-textured materials via self-organization processes or as templates for the growth of nanowires, nanorods, and nanodots [7]. Nanostructure engineering by energetic ion beams also includes patterning or ordering of the magnetic properties of ultrathin ferromagnetic films [8, 9] and the fabrication of nanodots by single ions [10].

3 Nanomaterials for Enhanced Radiation Performance

Understanding of radiation effects in nanomaterials is an urgent challenge, since it may hold the key to unlock the design of new materials for advanced nuclear energy systems. With increasing demands for energy, the projected growth in nuclear power will require the development of radiation-tolerant materials. To support the expanded use of nuclear power, a new understanding is required for nanoscale interactions at interfaces, microstructural evolution, and phase stability in nuclear materials under high radiation conditions.

Nanostructured materials include a broad class of materials in the form of thin films, coatings, powders, or bulk materials. Conventional materials have grains sizes ranging from microns to several millimeters. Nano-sized grains contain many fewer atoms. As the grain size decreases, there is a significant increase in the volume fraction of grain boundaries or interfaces, which strongly influences the chemical and physical properties of the material.

Materials based on nanoscale interfaces have demonstrated improved radiation resistance [11]. Nano-spaced interfaces may act as effective sinks for point defects, enhance defect cluster annihilation, and promote Frenkel pair recombination. By controlling the grain size down to the nanoscale, the density of interfacial grain boundaries can be significantly increased, and enhanced radiation tolerance in nanocomposites may be realized because of the short diffusion distances for defects to the nearest sink. This enables the rapid removal of the point defects, suppression of helium bubble nucleation, and retardation of the bubble growth at elevated temperatures. Utilization of nanoscale interfaces and second-phase features to attract migrating point defects, increase their recombination rates, and resist radiation damage accumulation has been reported as the underlying mechanisms for the enhanced radiation performance of nanostructured materials. Fine dispersions of carbide and phosphide precipitates in advanced austenitic stainless steel were found to substantially increase the incubation dose for swelling [12, 13]. A substantial enhancement in radiation-induced amorphization resistance has been reported in single-phased nanocrystalline $MgGa_2O_4$ with grain sizes ranging from 4 to 12 nm versus large-grained polycrystalline ceramic with an average grain size of 10 μm [14], demonstrating that interfacial regions act as sinks for point defects and their clusters produced under irradiation.

Nanocrystalline or nano-layered metallic systems with improved radiation properties have been reported both experimentally and theoretically. Enhanced radiation performance is due to grain boundaries and layer interfaces that act as effective sinks for defect recombination. For example, multilayer Cu/Nb composites with nanodimensional interlayer spacing exhibit excellent resistance to irradiation-induced structural changes [15, 16]. Nanolayered Cu/Nb composites with layer thickness of 2.5 nm have shown ultra-high strengths and enhanced radiation damage tolerance [16]. Molecular dynamics (MD) calculations indicate that grain boundaries act as interstitial sinks, and a vacancy-dominated defect structure is obtained after solidification of cascade events [17]. Radiation response of pre-existing nanopores at the grain boundaries in nanocrystalline Fe was also studied by MD simulation,

and both the void and grain boundaries act as effective sinks to self-interstitial atoms formed from nearby displacement cascades [18].

Materials containing a high concentration of nanoscale interfaces and secondary phase features may offer very high resistance to radiation damage accumulation. However, nanostructured materials are not inherently resistant to radiation damage, and the irradiation-enhanced amorphization in nanocrystalline solids is strongly materials dependent. Defect absorption alters interface structures, interfaces can become saturated with defects or segregants, and damaged interfaces can act as sources of defects instead of defect sinks. Radiation can also atomically mix interfaces and layers, or destroy the geometrical arrangement of layered morphologies, because the flow of radiation-induced point defects into an interface changes its atomic configuration. Interface degradation may ultimately determine material response under irradiation. It has been shown that small ZrO_2 crystallites with an average particle size less than ~ 3 nm embedded in nanostructured ZrO_2/SiO_2 composites amorphize at a low ion irradiation dose (0.9 displacements per atom, dpa); whereas much larger ZrO_2 crystallites do not amorphize following very high ion exposures (680 dpa) [19–21]. A theoretical model [22] suggested that the competition between irradiation-induced formation of point defects and the absorption of point defects by grain boundaries defines the radiation resistance of nanocrystalline materials.

An atomic-level understanding of radiation effects at and near interfaces, including grain boundaries and precipitate-matrix interfaces, is crucial for the fabrication of advanced nuclear materials that can withstand the extreme radiation environments in reactors, accelerators, and nuclear waste stored in geologic repositories. More experiments and theoretical efforts are required in order to clarify the damage mechanisms in nanostructured materials and to understand the effects of the size of crystalline nanoparticles, associated with increased surface energy on the damage production, accumulation, and annihilation.

4 Scientific Challenges

Much of the work on radiation effects in oxides has focused on atomic collision processes [23–27] in terms of defect formation, migration, recombination, and microstructure evolution. The effects of ionization (electronic excitations) [28–32] and the coupled effects of ionization and atomic dynamics on material behavior are not well understood and constitute a challenging area of research [33].

Energy deposition into the atomic structure occurs by kinetic energy transfer that leads directly to defect production through atomic displacements. Basic processes of defect production by ballistic collisions on the atomic structure are fairly well understood [24]. Energy transfer processes between atoms involve elastic collisions that are reasonably well described by a screened interatomic potential, and the entire energy cascade (temporal and spatial characteristics) can be simulated by computational approaches, such as empirical potential molecular dynamics. However,

very little is known about residual defect migration and aggregation processes after the ballistic phase of cascade formation and after the quench has ceased. The atomic dynamics of defect evolution between cascade events and from cascade accumulation cannot be accurately simulated using existing methods. Furthermore, the effects of mixed covalent and ionic bonding in many ceramic materials on the dynamics of these atomic processes are not understood. New interatomic potential models and improved codes are required to simulate atomic dynamics in these mixed bonding systems.

Electronic excitations induced by ionization energy losses in materials can cause the trapping of charge at defects, as well as local changes in bonding and free energy; these can lead to permanent bond breakage and the formation of defects in the atomic structure (solid state radiolysis), local changes in barriers to atomic diffusion, and phase instabilities. Under extreme ion and electron fluxes available in the laboratory today, some electronic excitations from ionization can have lifetimes on the order of the time between events. As a result, a steady-state concentration of electronic excitations can exist that has a significant effect on the evolution of microstructure, phase changes, and thermodynamic properties in materials; and these localized excitations can influence the complex interaction of materials with the environment. New computational and experimental approaches are needed to develop predictive models to describe such effects.

Coupled effects of ionization and atomic dynamics on materials behavior are even less understood; for example, the effects of simultaneous ionization-induced electronic excitations on ballistic defect production, defect migration, and damage accumulation processes. Currently, simulation and modeling of radiation effects in materials primarily focuses only on processes occurring in the atomic structure. There is little understanding of how ionizing radiation creates defects directly or affects the kinetics of atomic processes through enhanced defect diffusion or atomic mobility. Consequently, there is a great need for revolutionary advances in understanding and modeling of radiation effects that address ionization effects from charged-particle irradiation. By understanding and modeling the coupling of electronic and atomic dynamics over a range of irradiation conditions, it should be possible to predict coupled behavior under the low dose rate, long-term irradiation conditions, and provide the bridge to macroscopic observations and system level material performance models.

References

1. T. Kunert, and R. Schmidt, Excitations and fragmentation mechanisms in ion–fullerene collisions, Phys. Rev. Lett. 86 (2001) 5258–5261.
2. F. Banhart, and P. M. Ajayan, Carbon onions as nanoscopic pressure cells for diamond formation, Nature 382 (1996) 433–435.
3. W. Mickelson, S. Aloni, W. Q. Han, J. Cumings, and A. Zettl, Packing C60 in boron nitride nanotubes, Science 300 (2003) 467–469.
4. L. Sun, et al. Carbon nanotubes as high-pressure cylinders and nanoextruders, Science 312 (2006) 1199–1202.

5. J. Lian, L. M. Wang, S. X. Wang, J. Chen, L. A. Boatner, and R. C. Ewing, Nanoscale manipulation of pyrochlore: New nanocomposite ionic conductors, Phys. Rev. Lett. 87 (2001) 145901.

6. J. Lian, W. Zhou, Q. M. Wei, and L. M. Wang, Simultaneous formation of surface ripples and metallic nanodots induced by phase decomposition and focused ion beam patterning, Appl. Phys. Lett. 88 (2006) 093112.

7. U. Valbusa, C. Boragno, and F. B. de Mongeot, Nanostructuring surfaces by ion sputtering J. Phys. Condens. Matt. 14 (2002) 8153.

8. C. Chappert, et al., Planar patterned magnetic media obtained by ion irradiation, Science 280 (1998) 1919–1922.

9. H. Bernas, et al., Ordering intermetallic alloys by ion irradiation: A way to tailor magnetic media, Phys. Rev. Lett. 91 (2003) 077203.

10. E. Akcöltekin, et al., Creation of multiple nanodots by single ions, Nat. Nanotechnol. 2 (2007) 290–294.

11. K. Kaneko, Y. Furuya, M. Kikuchi, High-temperature strength of TiC-coated SUS316 stainless steel Fusion Eng. Des. 19 (1992) 293.

12. E. H. Lee, A. F. Rowcliffe, and L. K. Mansur, Precipitation and cavity formation in stainless steels during irradiation, J. Nucl. Mater. 103 and 104 (1981) 1475.

13. P. J. Maziasz, Void swelling resistance for PCA austenitic stainless steel under HFIR irradiation through microstructural control, J. Nucl. Mater. 200 (1993) 90.

14. T. D. Shen, S. Feng, M. Tang, J. A. Valdez, Y. Wang, and K. E. Sickafusa, Enhanced radiation tolerance in nanocrystalline $MgGa_2O_4$, Appl. Phys. Lett. 90 (2007) 263115.

15. T. Höchbauer, A. Misra, K. Hattar, and R. G. Hoagland, Influence of interfaces on the storage of ion-implanted He in multilayered metallic composites J. Appl. Phys. 98 (2005) 123516.

16. A. Misra1, M. J. Demkowicz, X. Zhang, and R. G. Hoagland, The radiation damage tolerance of ultra-high strength nanolayered composites, *JOM J. Miner. Metals Mater. Soc.* 59 (2007) 62–65.

17. M. T. Robinson, Basic physics of radiation damage production J. Nucl. Mater. 216 (1994) 1.

18. M. Samaras, W. Hoffelner, and M. Victoria, J. Nucl. Mater. 352 (2006) 50–56.

19. A. Meldrum, L. A. Boatner, and R. C. Ewing, Nanocrystalline zirconia can be amorphized by ion irradiation Phys. Rev. Lett. 88 (2002) 025503.

20. A. Meldrum, L. A. Boatner, and R. C. Ewing, Size effects in the irradiation-induced crystalline-to-amorphous transformation Nucl. Instrum. Methods Phys. Res. B 207 (2003) 28–35.

21. K. E. Sickafus, H. Matzke, T. Hartmann, K. Yasuda, J. A. Valdez, P. Chodak, M. Nastasi, and R. A. Verrall, Radiation damage effects in zirconia J. Nucl. Mater. 274 (1999) 66.

22. I. A. Ovid'Ko and A. G. Sheinerman, Irradiation-induced amorphization processes in nanocrystalline solids Appl. Phys. A 81 (2005) 1083.

23. S. Theppakuttai and S. Chen, Submicron ripple formation on glass surface upon laser-nanosphere interaction, J. Appl. Phys. 95 (2004) 5049.

24. W. J. Weber, R. C. Ewing, C. R. A. Catlow, T. Diaz de la Rubia, L. W. Hobbs, C. Kinoshita, H. J. Matzke, A. T. Motta, M. Nastasi, E. K. H. Salje, E. R. Vance, and S. J. Zinkle, Radiation effects in crystalline ceramics for the immobilization of high-level nuclear waste and plutonium J. Mater. Res. 13 (1998) 1434.

25. K. E. Sickafus, L. Minervini, R. W. Grimes, J. A. Valdez, M. Ishimaru, F. Li, K. J. McClellan, T. Hartmann, Radiation tolerance of complex oxides Science 289 (2000) 748.

26. Y. Zhang, W. J. Weber, V. Shutthanandan, R. Devanathan, S. Thevuthasan, G. Balakrishnan, D. M. Paul, Damage evolution on Sm and O sublattices in Au-implanted samarium titanate pyrochlore J. Appl. Phys. 95 (2004) 2866.

27. Y. Zhang, W. J. Weber, W. Jiang, A. Hallén, and G. Possnert, Damage evolution and recovery on both Si and C sublattices in Al-implanted 4H-SiC studied by Rutherford backscattering spectroscopy and nuclear reaction analysis J. Appl. Phys. 91 (2002) 6388.

28. Y. Zhang, J. Lian, C. M. Wang, W. Jiang, R. C. Ewing, and W. J. Weber, Ion-induced damage accumulation and electron-beam-enhanced recrystallization in $SrTiO_3$ Phys. Rev. B 72 (2005) 094112.
29. I.-T. Bae, Y. Zhang, W. J. Weber, M. Higuchi, and L. A. Giannuzzi, Electron-beam induced recrystallization in amorphous apatite Appl. Phys. Lett. 90 (2007) 021912.
30. H. Inui, H. Mori, T. Sakata, H. Fujita, Electron irradiation induced crystalline-to-amorphous transition in quartz single crystals. J Non-Cryst Solids 116 (1990) 1.
31. L. W. Hobbs and M. R. Pascucci, Radiolysis and defect structure in electron-irradiated α-quartz J. Phys. 41 (1980) C6-237.
32. W. L. Gong, L. M. Wang, R. C. Ewing, and J. Zhang, Electron-irradiation- and ion-beam-induced amorphization of coesite Phys. Rev. B 54 (1996) 3800.
33. Y. Zhang, In-Tae, Bae, W. J. Weber, Atomic collision and ionization effects in oxides, Nucl. Inst. Methods Phys. Res., B 266 (2008) 2828.

Box 2: Sputtering

Hans Henrik Andersen

When heavy energetic ions impinge on a solid target, atoms may be expelled from the target surface. This erosion process is called *sputtering*. Depending on the ion-target combination and the ion energy, the number of target atoms expelled at perpendicular incidence may vary from the detection limit of 10^{-5} to tens of thousands. The most basic parameter of sputtering is the *yield Y*, defined as the ratio between the number of incident particles and emitted atoms.

Next to ion implantation, sputtering is the most important mechanism for the use of ion beams in nanotechnology. The flux of sputtered atoms generated from a bombarded target may be used for thin film deposition. Conversely, the removal of target atoms, often called sputter etching, may be used for layer removal and for building of two- or three-dimensional structures; the latter in combination with masking and etching. Sputter etching may also be used for mask repair and for layer-by-layer removal, for depth profiling of major and trace elements, and for other analytical purposes. Some of these applications are treated in later chapters of the present book.

The experimental and theoretical literature covering the subject of sputtering over the last 150 years is enormous and no systematic reference to the original publications shall be attempted here. The reader will mainly be guided to review papers in the four volumes edited by Rainer Behrisch [1–4] and to two sets of symposia proceedings edited by Peter Sigmund [5, 6].

Sputtering at perpendicular incidence involves a reversal of momentum. For that reason alone, several scattering processes must be involved to expel an atom from the surface. Primary target recoils collide with other target atoms creating secondary recoils, etc. A so-called *collision cascade* develops. The flow of particles and energy within the collision cascade may be described by a nonlinear Bolzmann transport equation. This nonlinear integro-differential equation may be solved in a linear approximation to find yields and energy and angular distributions of the sputtered flux. In the fully developed cascade, the flux of moving recoils is isotropically

H.H. Andersen (✉)
Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, DK-2100 Copenhagen O, Denmark
e-mail: nimb@fys.ku.dk

distributed. Those recoils that reach the surface with an energy sufficiently large to overcome the surface binding energy are sputtered. The linearization of the Bolzmann equation implies that all collisions within the cascade are between moving atoms and atoms that are stationary in the target. That is, the cascade is 'thin' or 'rarified.' Before the cascade totally dies out, its energy has been distributed to all atoms within its boundary and they will all move, but this limit will mostly not be reached until their energies are too low to cause sputtering. Sigmund summarized his theory of the processes described above in 1981 [7], and up-to-date discussions of recent developments are given by Jakas [8] and Urbassek [9].

The main theoretical results are that the energy dependence of the yield is given as being proportional to the nuclear stopping $S_n(E)$, (see chapter "Basis of Ion Scattering in Nanoscale Materials"). $S_n(E)$ also describes the dependence on Z_1 (projectile) and Z_2 (target), except for a correction at large mass ratios M_2/M_1. Further, the yield is inversely proportional to a surface binding energy U, usually taken to be the sublimation energy. The validity of the theory is restricted to projectile energies sufficiently large to generate a fully developed collision cascade (i.e., some keV) and to ion-energy combinations where electronic stopping is not significant. The latter condition excludes very light projectiles at all energies and all projectiles at very high energies. Only for insulating targets may electronic effects be of importance at energies relevant for nanotechnological applications (but see also the chapter "Micro- and Nanoengineering with Ion Tracks"). For heavy projectiles and heavy targets, we generate very dense collision cascades, and the linear approximation will break down as described below.

The agreement between theory and experiment may be judged from the large data collection by Andersen and Bay [10]. This paper may also be used to estimate the yield for any given combination of ion, energy, and mono-component target material through utilization of the theoretical scaling rules stated above. For that purpose it might, however, be more convenient to use the empirical formulae of Matsunami et al. [11] together with their data collection. These tables have the virtue that they, in contrast to the straightforward application of the solutions to the transport theory applications, also cover the low energy threshold region. Generally, a yield estimate within a factor of two may be obtained. An alternative approach to estimate a given yield is to run a Monte Carlo simulation, e.g., in the form of a SRIM 2003 or more recent version simulation (http://www.srim.org). Again, estimates within a factor of two may be obtained, but for low fluences only.

For all the above estimates it should constantly be kept in mind that the numbers refer to monocomponent targets, to flat target surfaces, and to amorphous or fine-grained polycrystalline targets. Therefore, large deviations (factor of five) from the above estimates may occur, whenever sputtering is carried out to high fluences, except for self- and noble gas-sputtering, where the fluence dependence may usually be kept within the factor-of-two uncertainty mentioned above.

All the above estimates relate to perpendicular incidence of the bombarding particles. For oblique incidence angles, it still holds that the yield is proportional to the energy deposited in the uppermost layers of the target. A zero-order estimate is hence that the yield is inversely proportional to the cosine of the polar angle of

incidence. For polar angles larger than approximately 60°, reflection of projectiles from the first surface layers starts to have an influence and the yield eventually passes through a maximum to approach zero at grazing incidence. For data see Andersen and Bay [10].

The *energy distribution* of the sputtered species is in its simplest theoretical version given by a function that has its maximum at approximately $U/2$, where U is the surface binding energy of the target material, i.e., at a few eV. At higher energies the distribution falls off asymptotically slightly faster than E^{-2}; see Sigmund [7], Andersen [12], Hofer [13], Wucher [14], Urbassek [9], and Gnaser [15]. The energies above the maximum of the distribution are of importance for the quality of thin films fabricated by sputter deposition, as they render some mobility to the deposited atoms without resputtering or creation of damage.

The *angular distribution* has been studied less than the energy one, but is of larger importance for applications. It has been discussed in some of the references quoted for the energy distribution. Schematically, the distribution at different projectile energies looks as shown in Fig. 1. Below the energy region of the developed cascade, conservation of momentum causes backwards emission (perpendicular to the surface) to be rare. For fully developed cascades, where the flux within the bombarded target is isotropic, it could be expected that the angular distribution of the emitted particles would be cosine, but the outward scattering by the surface atoms neighbouring the sputtered atoms cause the distribution to be overcosine. Regular cosine distributions are found only in a transition region. For oblique incidence at lower energies, the peak in the distribution in the forward direction gains in importance and appears closer to the surface. The asymmetry diminishes with increasing energy. The implications of the discussed angular distributions should be considered carefully both for film deposition and for depth profiling purposes.

The *mass distribution (clustering)* of the sputtered atoms will not be discussed here, but see Wucher [14]. The *charge distribution* is of course of crucial importance for analytical purposes. For most surfaces and species only a small fraction of the
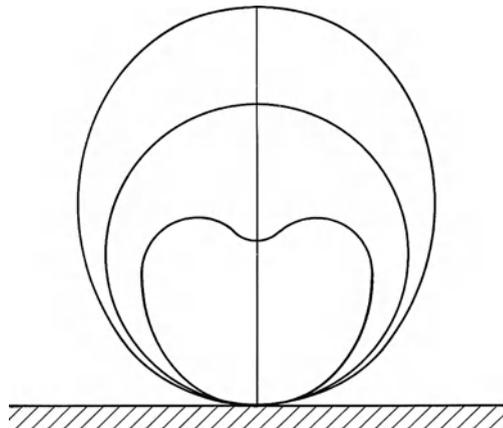

Fig. 1 Qualitative experimental angular distributions of sputtered particles for perpendicular incidence. With increasing projectile energy, the distributions vary from heart-shaped via cosine to over-cosine (From Andersen [12] with permission)

sputtered atoms are ionized (or negatively charged), but the charge state is strongly dependent on the chemical nature of the emitting surface. For a review see the paper of Wittmaack [16]. The dependence of the chemical state of the surface is utilized both in SIMS measurements and in sources for negative ions by bombarding with Cs ions. The easily ionized Cs atoms are neutralized on the surface of the sputter target, but do easily give off an electron. The fraction of negatively sputtered atoms or clusters is therefore strongly enhanced.

The *topography* of a sputtered surface develops as sputtering proceeds. In most cases the topography develops into more and more complicated features as a function of bombardment fluence through a combination of the dependence of yield on incidence angle and the occurrence of surface diffusion. In this way the development of surface features also depends on target temperature and bombardment flux. Differently oriented crystal grains may develop very differently and spectacular structures may be generated, as shown by Carter et al. [17], Aziz [18], and Chason and Chan [19]. The evolution of the topography may give rise to a fluence dependence of all the parameters discussed above.

All sputtering experiments, except self-sputtering, will as a function of fluence create binary targets. For all *multicomponent targets*, whether beam generated or created in other ways, *preferential sputtering* is the rule. Lighter target atoms are sputtered preferentially over heavier ones, which give rise to isotope effects, among others. The dominating effect is, however, that elements with a low surface binding energy are sputtered preferentially over strongly bound ones. Conservation of mass means of course that for an infinite and homogenous target the sputtered flux integrated over all emission angles will in the steady state have the target composition and one projectile species will be expelled per incoming one. The only exception to this rule occurs at such low energies that the self-sputtering yield is below one. In such a case, we are depositing a film. The time to reach equilibrium, and hence the modulation of depth profiled concentration distributions will depend on a large number of complicating mechanisms like Gibssian surface segregation, radiation enhanced (bulk) segregation, chemical and radiation enhanced diffusion, etc. The very complicated profiles that may result have been discussed qualitatively, but in some detail, by Andersen [20]; see also Betz and Wehner [21]. The mathematics of the action of a number of the processes above within the collision cascade has been discussed in a major paper by Sigmund and Lam [22].

One complicating feature of the sputtering of multicomponent targets will be highlighted here: the influence of the generated concentration profiles on the angular distribution composition of the sputtered flux. If the very surface layer is depleted in the concentration of some element relative to the next one or two layers (everything beyond that may be neglected), the surface-depleted element will have a higher than average concentration at perpendicular emission and be depleted at oblique emission angles, as illustrated in Fig. 2, showing emission from a Ni_5Pd target under Ar^+ and O^+ bombardment. Note first that Gibssian segregation causes the element with lower surface binding energy, and hence the preferentially sputtered one, to be enriched at the very surface compared to the layers immediately below. These layers below the very surface are then heavily depleted in the said element as compared

Fig. 2 Relative angular
distribution of Ni and Pd
sputtered from Ni_5Pd as a
function of polar angle for
$15\,keV\ Ar^+$ and $10\,keV\ O^+$
bombardment (From
Andersen [12] with
permission)
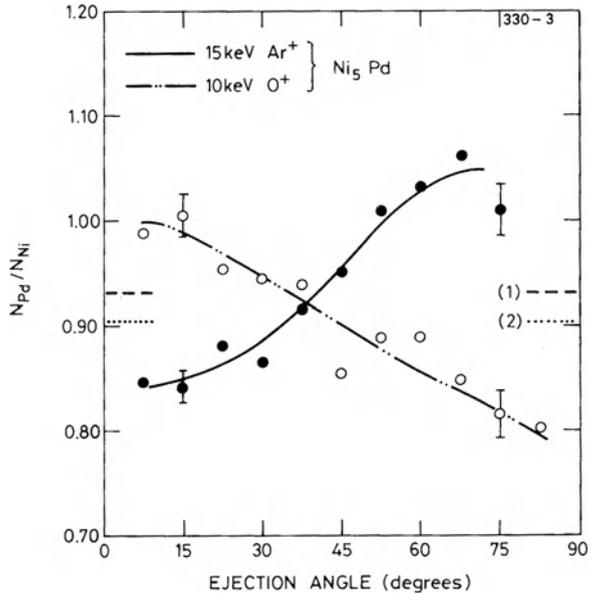
to the bulk. Note also, as seen from the figure, that the gradients may be reversed through bombardment with an element that preferentially reacts with one of the alloy components. These effects are of obvious importance, as well for deposition as for depth-profiling purposes.

Finally, we shall briefly mention the effects of nonlinearity. Within the theoretical treatment of the collision cascade, nonlinearity is defined as the breakdown of the linearization approximation. In practise, observable effects might be expected whenever the average energy of the atoms moving within the cascade is comparable to the bulk-binding energy of the target, as most recently discussed by Urbassek [9]. Experimentally, we shall define nonlinearity to occur whenever the sputtering yield of a cluster is different (in practice always larger) than the sum of the yield of the individual components impinging independently with the velocity of the cluster. Since the first observation of nonlinearity in sputtering (Andersen and Bay in 1974 [23]), a large body of data has been accumulated [24]. More recently, enhancement factors larger than 20 have been observed for metal targets bombarded by gold clusters [25]. Note, however, that yield enhancement may occur also for single-atom incidence of heavy atoms on heavy target surfaces. The condition is that we generate a very dense collision cascade.

The above-mentioned nonlinear effects have for some time been used technologically. Yamada et al. [26] developed a technique for surface smoothing, where bombardment with very large argon clusters at velocities much below the threshold for sputtering by individual atoms was used. The resulting sputtering through nonlinear effects had an angular distribution very much like the lower curve of Fig. 1, and the bombarded surfaces were made very smooth indeed. Bombardment with clusters

may also be used for depth profiling, as has been shown recently. The ability of large clusters to lift off intact complicated molecules from a surface, while causing little damage and mixing in deeper layers, would appear to become most useful in the future.

The present discussion has given a few rough guidelines for design of experimental setups and for interpretation of experimental results, but has mainly emphasized problems. For both cases the user is advised to study each individual case carefully.

References

1. Behrisch R (ed) (1981) Sputtering by ion bombardment. I. Springer, Berlin.
2. Behrisch R (ed) (1983) Sputtering by ion bombardment. II. Springer, Berlin.
3. Behrisch R and Wittmaack K (eds) (1991) Sputtering by ion bombardment. III. Springer, Berlin.
4. Behrisch R (ed) (2007) Sputtering by ion bombardment. IV. Springer, Berlin.
5. Sigmund P (ed) (1993) Fundamental processes in sputtering of atoms and molecules. Mat Fys Medd Dan Vid Selsk 43:1–666.
6. Sigmund P (ed) (2006) Ion-beam science – Solved and unsolved Problems. Mat Fys Medd Dan Vid Selsk 52:1–755.
7. Sigmund P (1981) Sputtering by ion bombardment: Theoretical concepts. In: Behrisch R (ed) Sputtering by ion bombardment I. Springer, Berlin, pp. 9–72.
8. Jakas MM (2004) Transport theories of sputtering. Phil Trans Roy Soc. A362:139–156.
9. Urbassek HM (2006) Sputter theory. In: Sigmund P (ed) Ion-beam science – Solved and unsolved problems. Mat Fys Medd Dan Vid Selsk 52:433–463.
10. Andersen HH and Bay HL (1981) Sputtering yield measurements. In: Behrisch R (ed) Sputtering by ion bombardment. Springer, Berlin, pp. 145–218.
11. Matsunami M, Yamamura Y, Itikawa Y, Itoh N, Kazumata Y, Miyagawa S, Morita K, Shimizu R, and Tawara H (1984) Energy dependence of the ion-induced sputtering yields of monatomic solids. At Data Nucl Data Tables 31:1–80.
12. Andersen HH (1988) Sputtering from atomic-collision cascades. Nucl Instr Methods B33:466–473.
13. Hofer WO (1991) Angular, energy and mass distribution of sputtered particles. In: Behrisch R (ed) Sputtering by ion bombardment III. Springer, Berlin, pp. 15–90.
14. Wucher A (2006) Sputtering experiments. In: Sigmund P (ed) Ion-beam science – Solved and unsolved problems. Mat Fys Medd Dan Vid Selsk 52:405–432.
15. Gnaser H (2007) Energy and angular distribution of sputtered species. In: Behrisch R (ed) Sputtering by ion bombardment IV. Springer, Berlin, pp. 1–90.
16. Wittmaack K (2006) A critical review of the electron-tunnelling model of secondary ion formation. In: Sigmund P (ed) Ion-beam science – Solved and unsolved problems. Mat Fys Medd Dan Vid Selsk 52:465–495.
17. Carter G, Navinsek B, and Whitton JL (1983) Heavy ion sputtering induced surface topography development. In: Behrisch R (ed) Sputtering by ion bombardment II. Springer, Berlin, pp. 231–269.
18. Aziz MJ (2006) Nanoscale morphology control using ion beams. In: Sigmund P (ed) Ion-beam science – Solved and unsolved problems. Mat Fys Medd Dan Vid Selsk 52: 187–206.
19. Chason E and Chan WL (2006) Ion-induced surface evolution in the linear instability regime: Continuum theory and kinetic Monte Carlo simulations. In Sigmund P (ed) Ion-beam science – Solved and unsolved problems. Mat Fys Medd Dan Vid Selsk 52: 207–225.

20. Andersen HH (1984) Ion-bombardment-induced composition changes in alloys and compounds. In: Williams JS and Poate JM (eds) Ion implantation and beam processing. Academic Press, Sydney, pp. 128–187.
21. Betz G, Wehner GK (1981) Sputtering of multicomponent materials. In: Behrisch R (ed) Sputtering by ion bombardment II, pp. 11–90.
22. Sigmund P and Lam NQ (1993) Alloy and isotope sputtering. In: Sigmund P (ed) Fundamental processes in sputtering of atoms and molecules. Mat Fys Medd Dan Vid Selsk 43:255–349.
23. Andersen HH and Bay HL (1974) Nonlinear effects in heavy-ion sputtering. J Appl Phys 45:953–954.
24. Andersen HH (1993) Nonlinear effects in collisional sputtering under cluster ion impact. In: Sigmund P (ed) Fundamental processes in sputtering of atoms and molecules. Mat Fys Medd Dan Vid Selsk 43:127–153.
25. Bouneau S, Brunelle A, Della-Negra S, Depauw J, Jacquet D, LeBeyec Y, Pautrat M, Fallavier M, Poizat JC, and Andersen HH (2002) Very large Au and Ag sputtering yields induced by keV to MeV Au_n clusters ($n = 1$–13). Phys Rev B 65:144106.1–144106.8.
26. Yamada I, Matsuo J, Toyoda N, Kirkpatrick A (2001) Materials processing by gas cluster ion beams. Mater Sci Eng R34:231–295.

Box 3: Ion Ranges

Yanwen Zhang

Ion range is one of the most important considerations in describing ion-solid interactions. When an energetic ion penetrates a solid, it undergoes electronic and nuclear stopping in the target. In the beginning of the slowing-down process at high energies, the ion is slowed down mainly by electronic stopping, and it moves almost in a straight path. As the ion energy decreases, the probability of collisions with nuclei increases, and the nuclear stopping finally dominates the slowing-down process. During the collision processes, target atoms (target recoils), which receive significant recoiling energies from the ion, will be removed from their lattice positions and produce a cascade of further collisions in the target.

The range, R, is determined by the rate of energy loss along the path of the ion,

$$R = \int_{E_0}^{0} \frac{1}{\mathrm{d}E/\mathrm{d}x} \mathrm{d}E, \tag{1}$$

where $\mathrm{d}E/\mathrm{d}x$ is the mean energy loss per unit path length. Because of the target structure, the nature of ion-solid interactions is statistical. Ions with the same energy incident at the same angle onto the sample surface do not necessarily come to rest at the same place. Some may undergo fewer collisions and travel further than others, as statistical fluctuations occur in the energy loss. Consequently, a statistical distribution of depths to which ions penetrate is observed. The ion distribution depends on several parameters, such as the ion mass, charge, and energy, the target mass, and the beam direction. The penetration distance along the incident ion axis is termed the *projected range*, R_p. Since the number of collisions required to bring ions to rest within the medium will vary with each ion, there will be a variation in the range, termed *straggling*, ΔR_p. A typical range distribution in an amorphous substrate for monoenergetic ions at a moderatedose is a symmetrical Gaussian shape in the first

Y. Zhang (✉)
Pacific Northwest National Laboratory, Richland, WA 99352, USA
e-mail: Yanwen.zhang@pnl.gov

approximation. Therefore, the distribution can be characterized by the two moments R_p and ΔR_p. The ion distribution, $C(x)$, for the implanted dose, D, can then be described by:

$$C(x) = \frac{D}{\sqrt{2\pi}\,\Delta R_p} \exp\left[\frac{-(x - R_p)^2}{2(\Delta R_p)^2}\right],$$ (2)

where x is the distance from the target surface, measured along the beam direction. At low or high ion energies, higher moments related to the *skewness* and the *kurtosis* are often necessary to describe the non-Gaussian ion profile. The *skewness* of the projected range distribution provides a measure of the asymmetry of the distribution. The *kurtosis* is a measure of the flatness of the distribution and is ~3 for most cases. At high doses, $\geq 10^{16}$ ions cm^{-2}, where the concentration of implanted atoms approaches a few atomic percent, other effects, such as sputter erosion of the surface, radiation-enhanced diffusion, or formation of tightly bound chemical phases, may alter or limit the ultimate attainable range and concentrations.

When ions enter the central part of a low-index axial crystal direction or planar channel in a crystal, electronic and nuclear stopping will be reduced considerably. The maximum range, R_{max}, of ions incident along main crystallographic directions in the crystal can exceed their range in an amorphous target by factors of 2–50. This process is termed *channeling*.

Computer simulation methods to calculate ion ranges in solids have been developed since the 1960s. The basic idea in them is to follow the movement of the ion in the medium by simulating the collisions with target nuclei. The electronic stopping power is usually taken into account as a frictional force slowing down the ion. Conventional methods used in computer simulations of the interactions of energetic atoms with solid targets are based on the binary collision approximation (BCA) [1–3]. In these methods, the movement of ions in the implanted sample is treated as a succession of individual collisions between the ions and the target atoms. While mainly a high-energy approximation, the BCA retains qualitative significance at low energies and, with proper formulation, gives useful quantitative information with good statistics in many situations.

The most well known BCA simulation program is TRIM (TRansport of Ions in Matter) [4, 5] developed by Ziegler, Biersack, and Littmark in 1985. The TRIM code simulates the slowing down and scattering of energetic ions in amorphous targets. In the TRIM simulations (http://www.srim.org/), the entire trajectories of large numbers of ions in an amorphous target are computed using a Monte Carlo method, where each projectile trajectory is statistically followed by randomly selecting a target atom, an impact parameter, and a distance (mean free path). Ion range and depth distribution of incident ions, collision cascades arising from the recoiled target atoms, displacements in the target, as well as angular and energy distributions of backscattered and transmitted ions can be followed. Nevertheless, a recent study [6] has shown that electronic stopping powers of slow heavy ions predicted by the TRIM code may be over-estimated by a factor of two for materials containing light element.

While the TRIM code is well accepted for quick estimation of ion ranges in amorphous materials, several BCA programs that take into account the crystal structure have been developed for various applications, such as the MARLOWE code and Crystal-TRIM.

The development of the MARLOWE code started in 1974 by Robinson and Torens [1, 2, 7]. Based on the BCA technique, the program considers crystalline target materials. The particle trajectories are constructed as series of binary encounters between the projectiles and the initially stationary target atoms. Elastic scattering is governed by one of several interatomic potentials, including multiple collisions in the calculations. The interactions, which bind atoms into crystals, are modeled by including binding energies between atoms and their original lattice sites, as well as binding to the entire crystal. Inelastic (electron excitation) effects are included in a low-energy approximation. Provision is made for users to supply alternative interatomic potential functions, inelastic energy-loss functions, initialization procedures, and analysis procedures. However, this calculation could be run only on a mainframe computer because of the tremendous computational effort.

The program Crystal-TRIM (http://www.fzd.de/pls/rois/) was developed based on the MARLOWE and TRIM codes. It simulates ion implantation into crystalline silicon, germanium, and diamond with up to 10 overlayers of arbitrary composition. Not only atomic ions but also molecular ions may be considered. Dynamic simulation of damage accumulation in crystalline substrates and the formation of amorphous layers are possible. The code can be used to calculate as-implanted range and damage distributions as function of depth. An efficient splitting algorithm is employed in order to enhance the statistical accuracy of the simulation results without a considerable increase in computing time. It is particularly useful if channeling tails are of interest.

References

1. M. T. Robinson and I. M. Torens, "Computer simulation of atomic-displacement cascades in solids in the binary-collision approximation," *Physical Review B*, 9(12), 5008–5024, 1974.
2. M. T. Robinson, "The binary collision approximation: Background and introduction," *Radiation Effects and Defects in Solids*, 130(1), 3–20, 1994.
3. G. Hobler and S. Selberherr, "Monte Carlo simulation of ion implantation into two- and three-dimensional structures," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 8(5), 450–459, 1989.
4. J. F. Ziegler, J. P. Biersack, and U. Littmark, In the Stopping and Range of Ions in Matter, Vol. 1, Pergamon, New York, 1985.
5. J. P. Biersack and L. G. Haggmark, "A Monte Carlo computer program for the transport of energetic ions in amorphous targets," *Nuclear Instruments and Methods*, 174, 257, 1980.
6. Y. Zhang, I.-T. Bae, K. Sun, C. Wang, M. Ishimaru, Z. Zhu, W. Jiang, and W.J. Weber, Damage Profile and Ion Distribution of Slow Heavy Ions in Compounds, J. Appl. Phys., 105 (2009) 104901.
7. M. T. Robinson, "Slowing-down time of energetic atoms in solids," *Physical Review B*, 40(16), 10717–10726, 1989.

Computer Simulation Methods for Defect Configurations and Nanoscale Structures

Fei Gao

1 Introduction

The selectively electrical doping of high-temperature electronic materials, such as SiC, GaAs, and GaN, using ion implantation, is an important technique that has been widely used in semiconductor industries. However, ion irradiation and the ion beam process create atomic-level defects, nano-defect structures and nano-amorphous domains, particular under ion-implantation to high dopant concentrations, the accumulation of which can cause amorphization, dopant diffusion, and dopant–defect interactions. These defects and nanoscale structures generated by ion–solid interaction prevent the electrical activation of the implanted dopant atoms and can affect the performance of semiconductor-based devices. In order to understand the mechanisms that control the defect generation, defect clustering, amorphization, and nanoscale structures during ion–solid interaction, there has been significant activity in modeling effects of energetic–beam interactions, such as those encountered in ion-beam processing or in neutron irradiation environments in metals [1–3], ceramics [4–6], and oxides [7, 8]. Investigations of ion-induced defects and nanoscale structures in semiconductors have been carried out by employing a wide range of analytical techniques (see reviews [9] and [10]), but a complete understanding of beam–solid interactions demands computer simulations to obtain insight and fundamental understanding of the complex atomic-level processes controlling the evolution of defects and nanoscale structures.

The microstructural changes and structural evolution under ion implantation or irradiation are inherently multiscale phenomena, which involve processes spanning a wide range of length and time scales. The successful development of understanding and predictive models for these multiscale phenomena requires integration of ab initio electronic structure calculations, molecular dynamics (MD) simulations, long-time scale dynamics, and kinetic Monte Carlo (KMC) methods

F. Gao (✉)

Pacific Northwest National Laboratory, MS K8-87, Richland, WA 99352, USA

e-mail: fei.gao@pnl.gov

over the relevant length and time scales to model and predict defect properties, defect generation and diffusion, defect accumulation, phase transformation, and microstructural evolution under ion–solid interaction.

Density functional theory (DFT) calculations require large computational efforts compared to the large-scale computer simulations normally applied to ion–solid interactions, such as MD and KMC simulations. However, ab initio DFT calculations can provide a set of accurate defect properties (formation, binding, and migration energies of point defects and small defect clusters). These results provide not only key input parameters for KMC models, but also guidelines for the development and fitting of semi-empirical potentials. During last decade, MD method has emerged as a main computational tool for studying defect properties, defect interaction, and diffusion, particularly in the form of clusters. In addition to defect cluster energetics, MD simulations of displacement cascade evolution have provided not only the mechanisms that control the defect generation, but also a database of primary damage production for the successful development of the multiscale modeling of microstructure evolution. The applications of MD simulations to more complex materials, such as SiC and GaAs, have until recently been more limited. With semi-empirical short-range potentials, the large MD cell containing up to 100 million atoms has been achieved for times approaching 100 ns; but the defect accumulation and nanostructures at the temperatures that defects become mobile, defect annealing, and second phase formation occur over seconds and microns. These studies are inaccessible to MD and must be modeled using either KMC or rate theory. The KMC method is similar to rate theory in which kinetic reaction-diffusion equations are solved, and can be used for upscaling atomic-level calculations to the macroscopic regime of experimental measurements. KMC solves the equations in fully three-dimensional space and thus captures all of the inherent spatial correlation and fluctuation. The KMC simulations require information on defect migration energies, binding energies, and interactions energies, as well as the initial defect configurations produced and the initial microstructural features of the materials. These parameters can be obtained from ab initio calculations, MD simulations and experiments.

2 Ab Initio Method

The term ab initio means from first principles, which does not mean that we are solving the Schrödinger equation exactly, but we are selecting a method that in principle can lead to a reasonable approximation to the solution of the Schrödinger equation. A wide range of ab initio methods have been employed in the calculations of defects and defect processes in recent years, particularly the Hartree-Fock method and density functional theory. Although density functional theory (DFT) has been around for many years, it is only recently (since 1990) that its application to defect engineering and nanoscale structures induced by ion implantation process has become feasible.

The density functional theory (DFT) treats the electron density as the central variable rather than the many-body wavefunction. This conceptual difference leads to a remarkable reduction in difficulty: the density is a function of three variables, i.e., the three Cartesian directions. The density functional theory was originally introduced by Hohenberg and Kohn [11] and Kohn and Sham [12], and allows us to reduce the problem of an interacting many-electron system to an effective single-electron problem. Writing the Euler-Lagrange equation in terms of the single-particle wave functions (ψ_i), the variation principle leads to an effective single-electron equation, well known as the Kohn-Sham (KS) equation [12]:

$$\left\{ -\frac{\hbar}{2m}\nabla^2 + V_{\text{eff}}(r) \right\} \psi_i(r) = \varepsilon_i \psi_i(r), \tag{1}$$

where ε_i is the KS eigenvalues, and V_{eff} is an external effective potential given by

$$V_{\text{eff}}(r) = V_{\text{ion}}(r) + V_{\text{H}}(r) + V_{\text{xc}}(r), \tag{2}$$

where V_{ion} is the Coulomb potential due to ions (or nuclei) plus possible other external fields, V_{H} the Hartree potential, and V_{xc} the exchange-correlation potential. To solve the KS equation requires precise knowledge of the exchange-correlation potential. It is the local density approximation (LDA) [12] which describes the exchange-correlation energy by taking the exchange-correlation energy per electron at a point r in the electron gas, to be equal to the exchange-correlation energy per electron in a homogeneous electron gas that has the same density as the electron gas at point r. The system of an infinite number of noninteracting motions in an effective potential within the DFT requires an infinite number of electronic weave functions. However, this issue can be further approximated using the concept of the periodicity of the crystal within the framework of Bloch's theory. Exploiting the lattice periodicity using Bloch's theorem has now led to the one-electron wavefunctions being expressed in terms of a Fourier expansion using plane waves as a basis set. Although plane waves are not the only possible basis set that can be used (for example, one could use atomic wavefunctions as a basis set), plane waves are perhaps more aesthetically appealing. Generally, the use of plane wave as a basis set leads to the cubic scaling of plane wave DFT calculations with system size [13, 14]. Accordingly, most efforts at achieving methods that scale linearly with system size have concentrated upon localized basis sets [15, 16].

For many years now, density functional theory has been used with great success to investigate the ground state properties of solids, both in the bulk form and at surfaces or interfaces. Particularly, the DFT method has been widely employed to calculate defect properties in a number of semiconductors (Si [17, 18], SiC [19, 20], GaAs [21, 22], and GaN [23, 24]) and pyrochlores [25, 26]. Figure 1 shows possible interstitial configurations in GaN, and similar configurations exist for N interstitials. At the T site, which is midway between two nonbonded Ga and N atoms along the c-axis, the interstitial has two nearest neighbors and six second-nearest neighbors (Ga tetrahedral–Ga_T or N tetrahedral–N_T). The octahedral interstitial,

Fig. 1 Schematic illustration
of **(a)** Ga octahedral
interstitial, Ga_O, and Ga
tetrahedral interstitial, Ga_T,
and **(b)** Ga split interstitials at
Ga site along $<11\bar{2}0>$
direction, Ga^+-Ga, and at N
site, Ga^+-N, where +
indicates the interstitial atom.
The large *dark spheres* show
Ga atoms and the large *light
gray spheres* the Ga
interstitials, and the *small
gray spheres* display N atoms

(a)   (b)
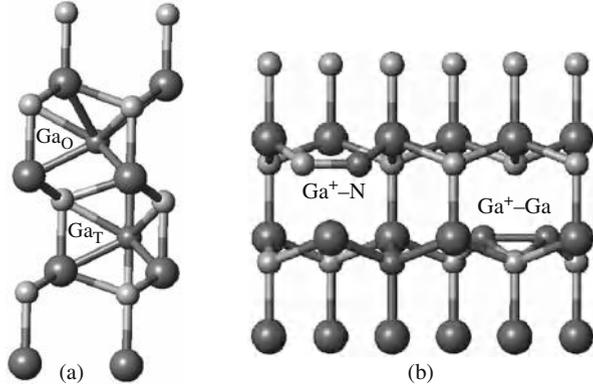
(Ga octahedral–Ga_O or N octahedral–N_O) is located at the center of a c-axis hexag-
onal channel between the Ga and N planes, and has three N and three Ga nearest
neighbors. The two possible configurations for Ga split interstitials are shown in
Fig. 1(b), all of which are along the $<11\bar{2}0>$ direction, where the Ga^+–N repre-
sents the split interstitial that forms with a Ga atom and is centered on the N site
and the Ga^+–Ga is the split interstitial centered on the Ga site. All the possible
defect configurations are relaxed using the DFT method, and the formation energies
are determined. The defect formation energies for the N-rich condition are given
in Table 1, along with those calculated by other ab initio methods [23, 24, 27–29]
for comparison. From Table 1, it can be seen that all calculations give similar for-
mation energies for the interstitials considered, but there are more diverse values
for the N vacancy, ranging from 1.2 to 5.0 eV. However, these calculations sug-
gest that the N interstitials, starting from all possible sites, eventually relax into
a N^+–N $<11\bar{2}0>$ split interstitial. In the case of Ga interstitials, the most stable
configuration is a Ga octahedral interstitial, but the energy difference among all

Table 1 Formation energy of defects, E_f, calculated by the DFT method for N-rich condition in
GaN, together with a full potential linear muffin-tin orbital method [27] for comparison, where
energies are given in eV

Defect	E_f [4]	E_f [23]	E_f [27]	E_f [28]	E_f [29]
V_{Ga}	6.14	6.8	6.3	6.3	6.7
V_N	4.81	1.2	4.6	5.0	2.8
Ga_N	9.59	6.8	10.5		
N_{Ga}	5.21	5.2	5.8		5.7
Ga_T	10.99				
Ga_O	10.56	4.1			
N_T	4.60				
N_O	4.58				
Ga^+–Ga $<11\bar{2}0>$	11.66				
Ga^+–N $<11\bar{2}0>$	10.99				
N^+–N $<11\bar{2}0>$	4.57	3.2			
N^+–Ga $<11\bar{2}0>$	4.61				

the interstitials is small. The Ga^+–Ga $<11\bar{2}0>$ split interstitial can bridge the gap between non-bonded Ga atoms, thereby leading to a chain of four Ga atoms along the $<11\bar{2}0>$ direction in GaN. Furthermore, electronic structures and bonding properties have been characterized using the Wannier function, which is an important approach to describe localized orbitals in materials, particularly for defects in solid-state phases. Figure 2 shows local Wannier orbitals of the N–N split, Ga–Ga split, and Ga octahedral interstitials [30]. Each N atom in the N–N split configuration has two σ orbitals that form covalent bonds between the N and Ga atoms, similar to those observed in perfect GaN. The unpaired electrons are localized on
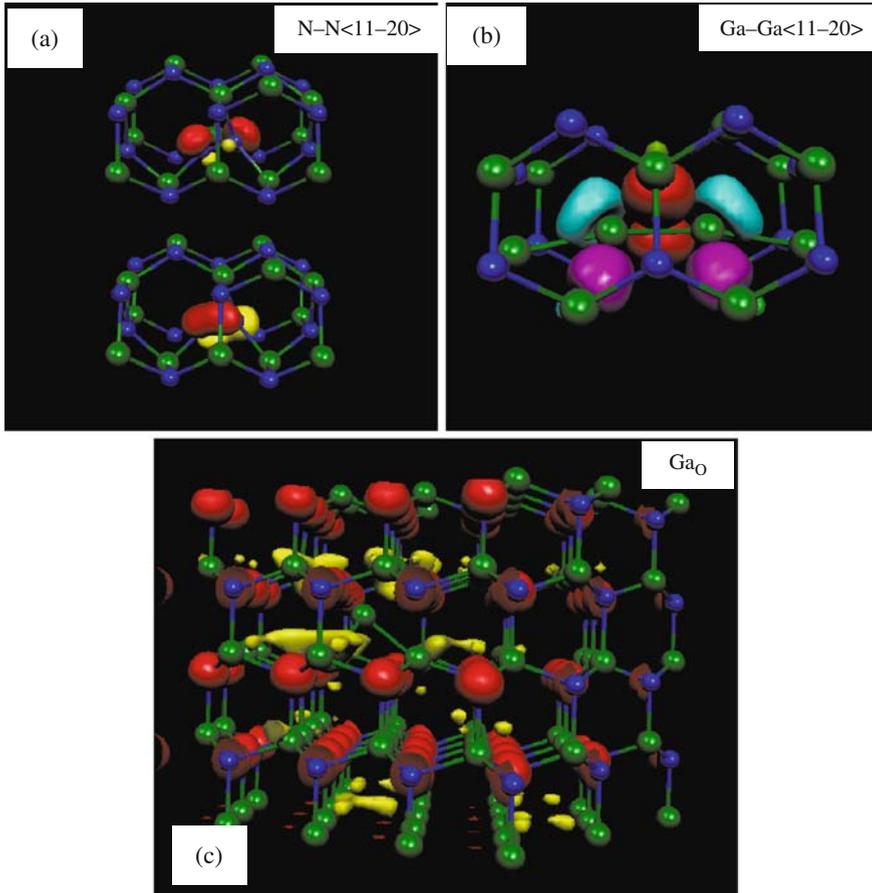
Fig. 2 Local Wannier orbitals of **(a)** the N^+–N $<11\bar{2}0>$ interstitial, **(b)** Ga_O interstitial, and **(c)** Ga^+–Ga $<11\bar{2}0>$ interstitial. The *top part* of **(a)** shows the two isolated pairs, each of which belongs to the N atoms in N^+–N $<11\bar{2}0>$ configuration, respectively, and the *bottom part* displays two π orbitals between N^+ and N atoms, where + indicates the interstitial. The *dark red* and *cyan* surfaces in **(b)** represent isolated pairs associated with the N and Ga atoms, respectively, but the σ orbitals are illustrated by the *purple* surfaces. The electronic orbitals shown in **(c)** are strongly delocalized, indicating that the extra electrons are in conduction band

the N atoms of the N–N split interstitials and form two lone pairs, one with each N atom, as shown in the top part of Fig. 2a. However, in the bottom part of Fig. 2a, the two π orbitals that form between the two N atoms, similar to those in an N_2 molecular, are clearly observed. It has been observed [30] that although the Ga antisite defect and N interstitials all lead to the formation of N–N split configurations, either parallel or perpendicular to the c axis, respectively, their local Wannier orbitals are remarkably different, which suggests that different electronic structures exit for these N–N split interstitial configurations. These differences are not easily revealed when considering only the configurations of the defects; thus, the Wannier function method provides a power technique to clearly describe the relevant electronic and bonding properties of defects. The local Wannier orbitals suggest that each Ga atom in the Ga^+–Ga split has only one σ orbital, forming a covalent bond with the nearest N atom, and also each Ga atom has a lone pair, as indicated by light gray. Furthermore, the unpaired electrons appear in the second nearest-neighbor N atoms, as marked by dark, which suggests a weak interaction between these atoms. The normally unbonded Ga atoms along the $<11\bar{2}0>$ direction form metallic-like bonds with the Ga^+–Ga split interstitial, with bond distances close to those in bulk Ga. These results suggest that the Ga interstitials can lead to the formation of Ga atomic wires in GaN, which could have significant implications for the application of GaN electronic devices, since such atomic-scale wires could have transport properties different from those in the bulk due to distinct quantum effects. Consequently, it may be possible to fabricate atomic quantum wires by introducing Ga interstitials during the ion implantation process. In the case of the Ga octahedral interstitial, the electron orbitals are greatly delocalized, and the extra electrons are in the conduction band, as shown in Fig. 2c. There are no covalent bonds formed between the Ga_O interstitial and the surrounded N atoms, indicating a weak interaction.

Several attempts have been carried out to calculate threshold displacement energies using ab initio molecular dynamics [31, 32], and these calculations demonstrate that such calculations are feasible, at least for covalent materials for which the vacancy-interstitial separation for the Frenkel pair is very small. However, the applications of ab initio MD method to nonequilibrium atomic-level defects and nanostructures, particularly to the primary state due to ion–solid interaction, are limited, and this may be due to the small size (100–1000 atoms) and short time ($\sim 5\,ps$) with which the ab initio method can deal with using the most powerful supercomputer. However, ab initio calculations provide the data sets for the development of atomic potentials for a much larger-scale molecular dynamics simulation. Although there are significant advancements in the development of semi-empirical potentials, most potentials employed are fitted to the various equilibrium properties of perfect crystals, which may not describe defect properties correctly. Recently, the force-matching method [33] has been developed to derive interatomic potentials based on the large amount of information that is obtained by first principle calculations. These potentials have been successfully applied to study surface properties [34], grain boundary structure [35], and nanoclusters [36] in transition metals. Also, the defect energies calculated using ab initio DFT [20], together with

various equilibrium properties in bulk SiC, have been employed to develop a new semi-empirical potential using a lattice relaxation fitting approach [37], which takes into account the chemical environment of the atoms, as well as their geometrical relationships with one another. This potential has been used to calculate the point defect properties in both 3C- [20] and 4H-SiC [38]. It was found that the potential derived from ab initio calculations describes defect behavior much better than the original Tersoff potential, and would provide for more accurate MD simulations of defect migration and cascade annealing processes.

3 Molecular Dynamics Simulations

Molecular dynamics method encompasses both cascade physics and the nucleation and growth processes of nano-amorphous domains and is currently an only viable approach for atomic-level investigation of the crystalline-amorphous (c-a) transition under irradiation. During the last decade, molecular dynamics methods have emerged as the main computational tool for studying defect properties, defect interaction, and diffusion, particularly in the form of nanoclusters. Some of the early studies of ion–solid interaction include the work of Gibson at al. [39], namely replacement collision sequences in copper, and that of Beeler and Besco [40]. However, the major boom in MD simulations began in the 1980s, when both computer capacity and atomic potential had advanced sufficiently so that it became possible to realistically display ion–solid simulations and to monitor the evolution of thousands of atoms for picosecond timescales. The MD simulations of Diaz de la Rubia et al. in 1987 [41] used large enough cells to fully contain the cascade, and the good agreement obtained with experimental mixing coefficients proved that the basic features of Brinkman's picture are indeed valid. At about the same time, several MD studies shed light on heat spike and low-energy sputtering (see, e.g., [42–45]), and in the early 1990s, numerous studies examined defect and dislocation production in collision cascades in metals and semiconductors.

A molecular dynamics simulation consists of the numerical solution of classical equations of motion, which may be written as

$$m_i \ddot{r}_i = f_i, \qquad f_i = -\frac{\partial V}{\partial r_i}, \tag{3}$$

where f_i is the force acting on atoms that can be derived from the potential energy of a system V, and r_i represents the atomic coordinates. There are various methods to solve Newton's equation of motion, including the Verlet algorithm [46], the velocity Verlet algorithm [47], and the predict-correct method [48]. Smith summarizes the present state of the field [49]. Even for 100^3 atoms, the number of surface atoms is about 6% of the total atoms, and this means that periodic boundary conditions need to be used in a displacement cascade study, unless surface effects are of particular interest. The size of the MD block of atoms is carefully chosen to avoid overlap of the cascade with itself by virtue of the periodicity. Of course, it is important to note

that special attention must be paid to the case where the potential range is not short, particularly for charged and dipolar systems. In general, a displacement cascade can be started by choosing a primary knock-on atom (PKA) and setting its energy and direction of motion to chosen values. The subsequent collision, displacements, and recombinations of atoms with vacant sites, and the formation of disordered domains or nanostructures, can be simulated in full detail, so that the complete structure and evolution of the cascade can be followed. It should be noted that most MD simulations of displacement cascades use a variable timestep, and actual timestep can be as small as 10^{-17} s during collisional phase. As soon as the extreme collisions are over, a much longer timestep is adequate, i.e., $10^{-15} - 10^{-14}$ s.

As has been reviewed elsewhere [50, 51], MD computer simulation has confirmed early ideas based on binary-collision concepts that a cascade exhibits two main stages as it evolves with time. The first is a ballistic or collision phase lasting a few tenths of a ps, during which the energy of the PKA is distributed by multiple collisions among many atoms, with the result that they leave their lattice sites. This creates a central disordered core surrounded by regions of crystal displaced outwards. In the second stage, a thermal-spike phase lasting several ps, the kinetic and potential components of the crystal energy attain equilibrium with each other. During this phase, the majority of the displaced atoms in the outer regions return by athermal relaxation to lattice sites, resulting in the formation of defects and defect clusters. Figure 3 shows the final defect nanostructure of a typical cascade of 20 keV in zirconium [52]. A feature that is apparent from the defect arrangements in Fig. 3 is that some large nanoclusters of vacancies and of interstitials are formed. The tendency of defects to cluster within their parent cascade is important, because it affects the way they behave and contribute to microstructure evolution. The detailed analysis reveals that most interstitial clusters have the form of a dislocation loop with a Burgers vector of

Fig. 3 Computer plot showing the final state of nano defects of a 20 keV cascade in Zr, where *white spheres* represent interstitials and *dark spheres* indicate vacancies
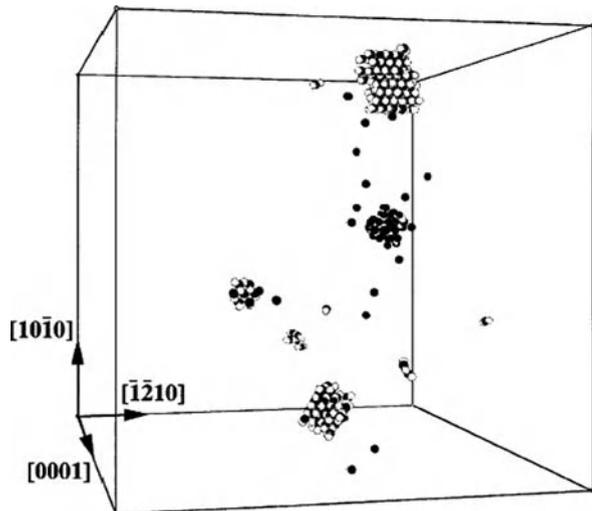
1/3 <11–20>, and can glide back and forth by one-dimensional migration along the direction of the Burgers vector. The collapse of the large vacancy cluster at the center of the cascade is observed, resulting in the formation of a vacancy dislocation loop on the prism-plane.

To further demonstrate the formation of nanoscale clusters and amorphous domains, MD simulation of Si displacement cascades with a modified Tersoff potential provides an excellent example for comparing nanoclustering behavior for Si and Au recoils in β-SiC [5, 53]. The primary damage states for 50 keV cascades from Si and Au recoil atoms are shown in Figs. 4(a) and (b), respectively, where only interstitials and antisite defects are included. Figure 4 shows only the part of the MD block that contains the Si or Au cascade. There are multiple branches along a 50 keV Si PKA path, which results in the formation of a very dispersed cascade that has distinct subcascades separated from each other and interspersed with isolated point defects. Most numerous defects are single interstitials and vacancies. The numbers of defects in clusters are few, and the size of the clusters is very small. These features are in agreement with the simulations of 10 and 50 keV Si recoil cascades [53, 54] under similar conditions. In contrast, the defect structures and cascade morphology produced by 10 and 50 keV Au PKAs are significantly different from those generated by Si PKAs, as shown in Fig. 4(b) for a 50 keV Au PKA. Although defect generation is highly dispersed within the cascade, several localized nanoscale disordered domains are created along the path of an Au PKA, as indicated by circles in Fig. 4(b). These regions have a very high defect concentration and consist of interstitials, vacancies (not shown), and antisite defects. This clearly demonstrates that large disordered domains can be created in SiC by an energetic heavy ion. While only one large disordered region can be formed in a 10 keV Au cascade, several nano-disordered domains are generated in a 50 keV cascade, as indicated by circles in Fig. 4(b).

To directly compare with experimental results, structural image simulation methods were applied to investigate the correlation between the known atomic arrangements of nano-disordered domains and simulated high-resolution transmission electron microscopy (HRTEM) images. The HRTEM image simulations were performed using the NCEM software [55], which uses a multislice technique based on a Fourier transformation. An area of the MD cell that contains the disordered domain, shown in Fig. 4(b), was cut into a small crystal containing about 3075 atoms, and the atomic arrangement is shown in Fig. 5(a). The irradiation-generated disordered domain can be seen in the center of the crystal. The simulated HRTEM image in Fig. 5(b) clearly shows the formation of a nanoscale-disordered domain within the SiC substrate produced directly within the Au track. In the disordered region, there is significant distortion and twisting of the atomic structure, and the image features are dominated by mottled contrast, similar to an amorphous structure. The interface between the disordered region and the crystal substrate is fairly well defined, and the bonds at the interface are highly twisted. The shorter bond length at the interface compared to the perfect crystal, and the misfit lattice structure, are evidence of lattice strain at the interface.

Fig. 4 Primary damage states
in the 3C–SiC MD cells:
(a) The displacement cascade
generated by a 50 keV Si
PKA and **(b)** the
displacement cascades
generated by a 50 keV Au
PKA, where only interstitials
and antisite defects are shown
for simplicity. The *large* and
small dark spheres represent
Si interstitials and antisites,
respectively, and *large* and
small gray spheres indicate C
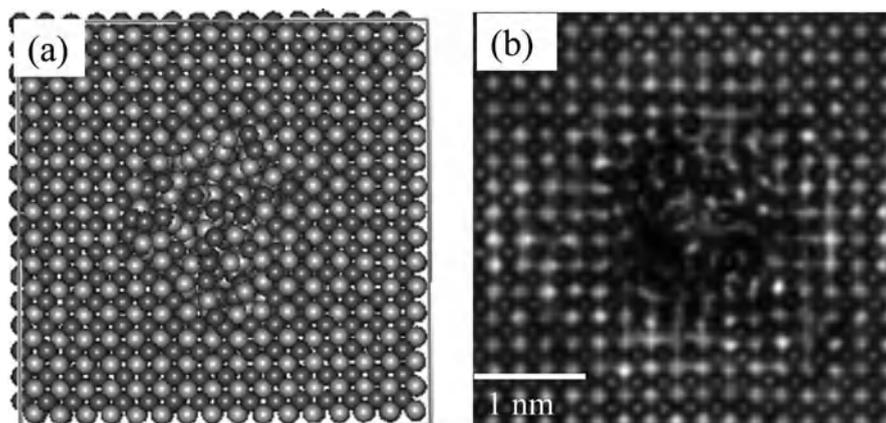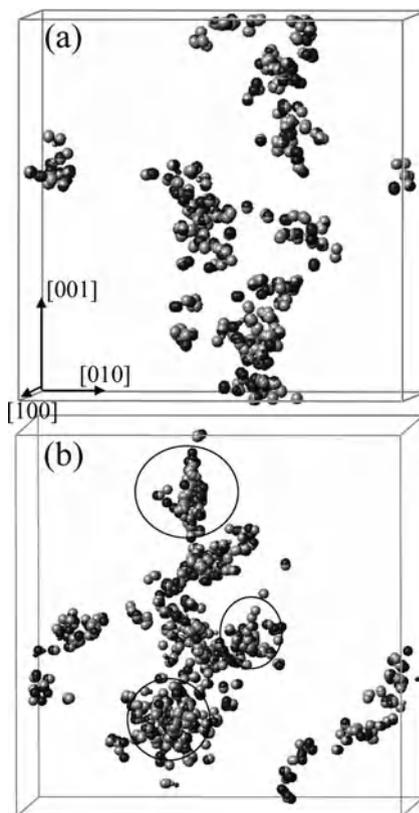interstitials and antisites,
respectively

Fig. 5 **(a)** Atomic plot of the part of MD cell that contains the disordered domain shown in Fig. 1
(*Light gray spheres* represent C atoms and *dark gray spheres* indicate Si atoms), and the corre-
sponding HRTEM image is shown in **(b)**, where the *larger spots* are Si and the *smaller spots* are C

   The damage structure shown in Fig. 5(b) is consistent with the interpretation of experimental data on damage accumulation and recovery in SiC irradiated with Si^+ and Au^{2+} ions [56]. The higher disordering rate and the residual disorder after thermal annealing observed for Au irradiation of SiC are associated with a higher probability for in-cascade formation of nanoscale disordered or amorphous domains. At present, there is no HRTEM confirmation of such large disordered cluster formation in SiC, due to the constraints of the experimental conditions.

   The simulation results have been used as input parameters in an up-scale analytical analysis, namely a direct-impact/defect-stimulated model [57] in which both amorphous nuclei (clusters) and defects can stimulate amorphization. Based on this model, the amorphous fraction, f_a, is given by the expression

$$f_a = 1 - (\sigma_a + \sigma_{ds})/\{\sigma_{ds} + \sigma_a \exp(\sigma_a + \sigma_{ds})D]\}, \qquad (4)$$

where σ_a is the cross-section for direct-impact amorphization, σ_{ds} is the cross-section for defect-stimulated amorphization, and D is the dose. In previous studies of SiC irradiated with either $2\,MeV\ Au^{2+}$ [58] or $550\,keV\ Si^+$ [59] ions at low temperatures, the amorphous fraction at the damage peak for each ion is approximated by the measured disorder on the Si sublattice, which is shown in Fig. 6 as a function of dose in displacements per atom (dpa). These energies were selected based on experimental constraints. In Fig. 6, full amorphization in SiC corresponds to a relative disorder of 1.0, and the solid lines are fits to the data using Eq. (4). The cross-sections can be determined from the model fits to the defect fraction of clustering obtained by MD simulations, where the ratio of σ_{ds}/σ_a describes the in-cascade ratio of isolated or freely migrating defects to the number of atoms in amorphous clusters. The analysis of cascade geometry in the MD simulations shows that the mixing feature of defects in a local region, which includes interstitials and antisite defects, can be identified only when a cluster contains four interstitials or more. It is, therefore, reasonable to consider a cluster containing four interstitials as a minimum amorphous cluster. The ratio of the defect-stimulated cross-section to
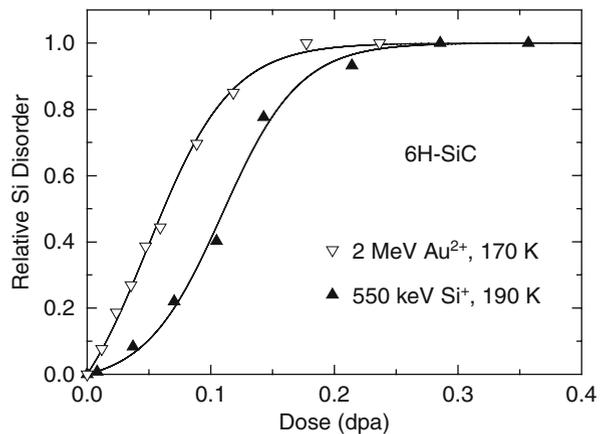
Fig. 6 Relative disorder on the Si sublattice as a function of dose (dpa) at the damage peak in 6H-SiC irradiated by $2.0\,MeV\ Au^{2+}$ at 170 K and $550\,keV\ Si^+$ at 190 K. *Solid lines* are the data fits using Eq. (4) in the text

the direct-impact cross-section has been calculated using the data in Fig. 6. Furthermore, the ratio of σ_{ds}/σ_a (Si) to σ_{ds}/σ_a (Au) represents the relative ratio of isolated interstitial formation to in-cascade clustering for Si and Au PKAs. This ratio is 7.31 from the MD simulations and 7.65 based on the experimental data and model fit. The excellent agreement of the relative ratio in-cascade clustering to isolated interstitial formation between MD simulations and the model fit to the experimental data suggests that the higher disordering rate for Au ions is associated with a higher probability for direct-impact amorphization or cluster formation, which increases the effectiveness of amorphization.

One of important applications of the MD method is to simulate the nanostructural features induced by ion–solid interaction due to multiple overlap of collision cascades in cubic (3C) SiC. While experimental studies of SiC have provided much data and some insights into irradiation-induced amorphization, the atomic-level and nanoscale processes controlling the crystalline-to-amorphization (c-a) transition in SiC are not well understood. Developing a fundamental understanding of the amorphization processes and growth of nano-amorphous domains represents an important scientific and technological challenge. In this simulation [60], the first cascade was initiated by imparting a kinetic energy of 10 keV to a Si atom chosen randomly near the top of the cell. The damage state was allowed to evolve for about 10 ps, after which it was equilibrated for another 10 ps to control temperature. The second and subsequent cascades were initiated using the same procedure with random direction and position. A total of 140 cascades were simulated, which corresponds to a dose level of 0.35 displacements per atom (dpa), or 27.5 eV per atom. The damage states in the central part of the MD cell for doses of 0.13 (52 cascades) and 0.28 dpa (112 cascades) are shown in Fig. 7(a) and (b), respectively, where only displaced atoms and antisite defects are shown. At low doses, the dominant defects are single interstitials and mono-vacancies, but even at a low dose of 0.05 dpa some small clusters are nucleated. The concentration of small clusters increases with increasing dose and the clusters are distributed randomly inside the MD block. The coalescence of small clusters due to high strain can occur at higher dose, forming nano-size clusters or nano-amorphous domains, as indicated by the circles in Fig. 7(a). These nano-amorphous domains consist of interstitials and antisite defects, which suggests that both defect species may play an important role in amorphization of SiC. Based on several different analyses, it appears that a completely amorphous state is reached at a dose of about 0.28 dpa, as illustrated in Fig. 7(b). The accumulation of disorder in 3C-SiC has been studied experimentally by ion implantation using in situ ion-channeling measurements [61] and HRTEM [62]. These experimental results have shown that point defect formation dominates at low dose levels (< 0.04 dpa). The present simulations are consistent with the experimental observations, which suggests that the dominant mechanisms for irradiation-induced amorphization in SiC are the homogeneous production of small clusters (amorphous nuclei) within individual cascades, and the stimulated growth and coalescence of these clusters by both point defects and small clusters that are produced during continuous cascade overlap.
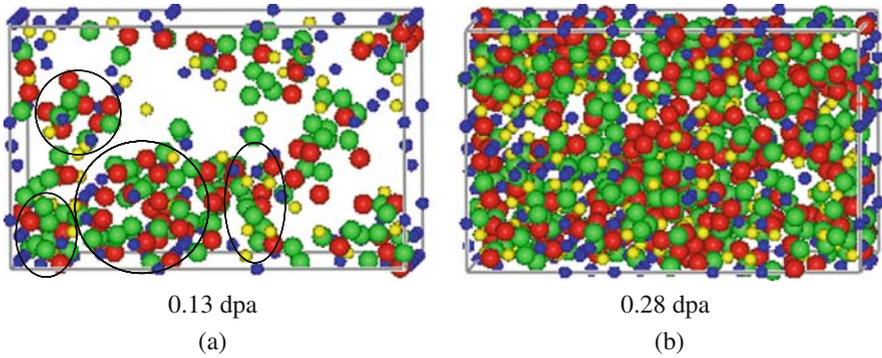
0.13 dpa   0.28 dpa

(a)   (b)

Fig. 7 Computer plots of damage state in the 3C–SiC MD cell at different damage levels: **(a)** 0.13 dpa and **(b)** 0.28 dpa, where *large* and *small dark spheres* represent Si interstitials and antisites, respectively, and *large* and *small gray spheres* indicate C interstitials and antisites, respectively

In order to study the correlation between damage accumulation during cascade overlap generated by MD simulation and the microstructural features observed experimentally, HRTEM images of cascade overlap states have been simulated. The HRTEM images simulated for damage states from cascade overlap and corresponding to doses of 0.13 and 0.28 dpa are shown in Fig. 8(a) and (b), respectively. In general, the present image simulations are consistent with the experimental HRTEM images obtained previously in irradiated 3C-SiC at similar damage states [62]. At an intermediate dose of 0.13 dpa (Fig. 8(a)), lattice stacking or long-range order is still present in some regions; however, there is greater distortion and twisting of the lattice fringes. These image features are dominated by the overlap of mottled and

0.13 dpa   0.28 dpa

(c)   (d)
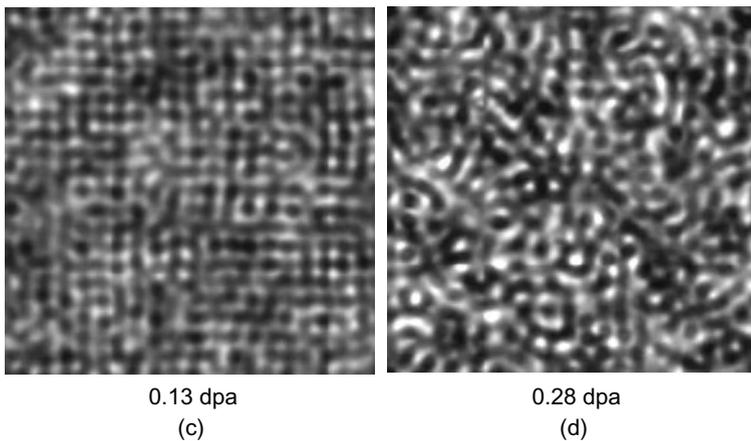
Fig. 8 The corresponding HRTEM image simulations of the 3C–SiC MD cell at the same damage levels: **(a)** 0.13 dpa and **(b)** 0.28 dpa, as shown in Fig. 7

lattice stacking contrasts that indicate the growth of nano-amorphous clusters or domains and the onset of a continuous amorphous network. The reasonable agreement between the MD image simulations and the previous experimental observations [62] at similar damage states is supportive of the cluster coalescence process in forming larger amorphous domains that was proposed above. After irradiation to a dose of 0.28 dpa, which corresponds to a nearly fully amorphous state, only mottled contrast is observed without any evidence for residual crystallinity. However, these MD simulations provide direct atomic-level insights into the interpretation of experimental HRTEM observations of damage accumulation and amorphization processes in SiC.

4 Kinetic Monte Carlo Simulations

The KMC method is similar to rate theory in which kinetic reaction-diffusion equations are solved, and can be used for upscaling atomic-level calculations to the macroscopic regime of experimental measurements. KMC solves the equations in fully three-dimensional space and thus captures all of the inherent spatial correlation and fluctuation. KMC simulation methods are generally employed to extend the spatial and time regimes of the MD simulations. While MD methods can follow the evolution of the damage and defect interactions for nanoseconds, KMC can follow the evolution of implantation damage, cascade annealing, and defect interactions over longer time and larger length scales. In KMC simulations, the underlying crystal structure is decorated with mobile and immobile lattice defects (such as interstitials, vacancies, implanted atoms, and substitutional defects) and nanoclusters, and the simulations follow the migration and interaction of these defects and nanoclusters on the crystal lattice. Each point defect or nano-defect cluster is associated with a lattice site of the materials, and the disposition of both mobile and immobile clusters is followed. A mobile defect or defect cluster jumps from one lattice site to an adjacent lattice site in a random direction when it is chosen by random selection from the distribution of all mobile defects and defect clusters. The distribution is weighted by the relative jump probabilities of the different defect species (size and type). The jump probabilities can be calculated from the equation

$$\sigma(j, n, T) = \sigma_0 \exp(-E_a/\kappa T), \tag{5}$$

where j is the defect type, n is the defect size, T is the temperature, and E_a is the activation energy for the corresponding defect. Generally, the dissociation energies are determined from the binding energies of the cluster. In addition to the spatial distribution of defect and defect clusters created by MD simulation (primary damage state), the KMC method requires a number of physical parameters, and these parameters can be determined using ab initio calculations and molecular dynamics simulations.

The KMC method has been successfully applied to defect annealing [63, 64] and accumulation simulations in metals [65, 66] and semiconductors [67] under ion–solid interaction or irradiation. These simulations describe how the defect distribution and accumulation evolve by diffusive processes, thereby providing the ability to perform atomic level studies of the defect kinetics and microstructure evolution over relevant length and time scales. One of important applications of the KMC method is to study nanoscale void lattice formation during irradiation [68, 69]. These simulations are based on the physical properties of defect clusters revealed by molecular dynamic simulations, i.e., glissile clusters of self-interstitial crowdions are formed directly from cascades, as seen in Fig. 3, and they migrate one-dimensionally along close-pack directions with extremely high activation energies. Starting from a random distribution of voids, the creation of stable void lattices is simulated by a KMC model in which vacancies migrate three-dimensionally and self-interstitial atom (SIA) clusters migrate one-dimensionally, interrupted by directional changes. Also, these simulations attempt to determine the microstructure conditions and nanocluster kinetics under which void lattices can form and to relate this information to experimental measurements of the spacing of void lattices and the void size. In this way, void lattice experiments may provide a direct link, through the KMC modeling, to characteristics of the primary defect production.

An advanced lattice kinetic Monte Carlo (LKMC) has been recently developed by Otto and Hobler [70], which couples between MD and LKMC for damage accumulation. Using this scheme, the dynamic annealing behavior of implantation damage for heavy and light ions has been studied in Si. This new approach has been further extended to investigate the recovery and clustering of defects during annealing of a single 10 keV cascade in cubic silicon carbide [71]. The defects initially created in a MD collision cascade are transferred to the KLMC model, and then are annealed isothermally from 100 to 1000 K in the KLMC model in a much larger cell and longer time. The defect states in the KLMC model after annealing at 300, 600, and 1000 K are shown in Fig. 9, along with the results for the initial cascade at 100 K (Fig. 9a). The number of surviving defects decreases with increasing temperature. After annealing up to a temperature of 1000 K, the damage cascade still retains much of its original shape, but with fewer defects. For example, the total numbers of point defects are about 178, 150, 140, and 124, respectively, at 100, 300, 600, and 1000 K. None of the vacancy (V) *clusters* or interstitial (I) *clusters* is observed to form at any of the annealing temperatures. The number of $V + I$ *complex* also does not change with annealing temperature. The number of A(antisite) $+ V + I$ *complex* increases slightly at high temperatures. The close Frenkel pairs play the most important role in recovery at temperatures below 300 K. The migration of isolated interstitials and point defect clustering becomes important at higher temperatures. Two distinct recovery stages for close Frenkel pairs are observed at about 200 and 550 K, and the growth of complex clusters is observed above 400 K. These simulation results are in good agreement with available experimental results [72, 73]. However, it is realized that the transfer of defect configurations between MD and MC is important for future multiple-cascade defect accumulation simulations.
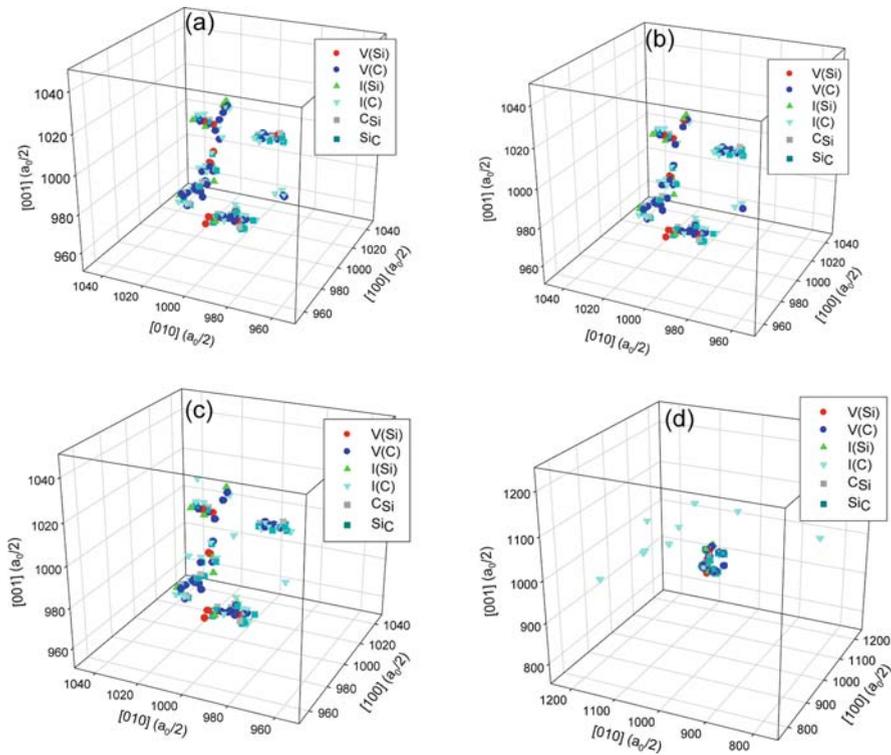
Fig. 9 Snapshots of defect survival after annealing at different temperatures in KLMC. **(a)**, **(b)**, **(c)**, and **(d)** are at temperatures of 0, 300, 600, and 1000 K, respectively

5 Other Methods to Extend Time Scale

As described above, molecular dynamics (MD) simulation technique, in which the classical equations of motion are integrated for a system of atoms, is a commonly used tool in chemistry, physics, and materials science. However, for processes occurring on time scales longer than nanoseconds, direct MD is computationally unfeasible. If the dynamics are characterized by "infrequent events," in which the system makes an occasional transition from one potential basin to another, rate constants can be determined from transition state theory, without ever performing dynamics, provided that the relevant saddle points can be found. However, the events that occur in real systems are often complicated and cannot be predicted in advance. How to treat these realistic systems, where the events are not known in advance, and occur on time scales inaccessible to MD, has been a long-standing problem. The key concept underpinning the *accelerated molecular dynamics methods* is, in a phrase, that trajectories are smarter than we are. While previous approaches for reaching long time scales have involved applying physical or chemical intuition to specify the states the system might evolve to, in fact the system often does something

unexpected, something outside our intuition. A way around this problem is to simply let the trajectory find the escape path for us. Each of the three accelerated dynamics methods described here is built on this foundation.

One of the most advanced techniques to extend the time scale of MD method is temperature-accelerated dynamics (TAD) [74]. This approach is designed to be easier to implement than hyperdynamics (no bias potential is required), at the cost of an additional approximation: harmonic transition state theory. By evolving the system at high temperature, transitions occur more rapidly, but not necessarily in the correct order. The TAD procedure filters out all the incorrect events, retaining the correct transitions and their transition times at the desired temperature. The migration of nanoclusters created by collision cascade in MgO has been studied using TAD simulations [75], with the time scale extended to a few tenths ns. While vacancies are essentially immobile, interstitials aggregate into clusters with surprising properties, i.e., an interstitial cluster with a size of four is immobile, but an impinging di-interstitial cluster can create a metastable cluster (six interstitials) that diffuses on a nanosecond time scale but is stable for years. These simulations well demonstrate how the accelerated molecular dynamics method can be applied to study ion–solid interaction in much longer time-scales.

Another interesting method is the dimer method [76]. The dimer method involves working with two atomic images of the system, and each of the images is displaced a small distance from their common midpoint. Searching for a saddle point involves rotating and translating the dimer. Each time the dimer is displaced, it is rotated so as to minimize the dimer energy, which is equivalent to finding the lowest curvature mode at the dimer midpoint. A saddle point search consists of finding a maximum along the lowest curvature mode, but a minimum along all other modes. To accomplish this, the dimer needs to be translated and moved uphill on the potential surface after each rotation. In the initial calculation, the dimer is minimized along a line defined by the initial force, and is then moved a small distance along the line to calculate the derivative of the effective force. Newton's method is used to estimate the zero in the effective force along the line and the dimer is then moved to that point. After each translation, the dimer is reoriented and moved along a direction conjugate to the previous line minimization. Also, Newton's method, along with quadratic approximation, is used to estimate the angle required to bring the rotational force to zero. Once a saddle point has been found, a normal minimization is performed from the saddle point to find the final configuration of the dimer. Possible transition states of defects in SiC [77] and α-Fe have been studied using the dimer method. It is found that a C interstitial can migrate from one site to another by multiple mechanisms, and the lowest energy barrier for its migration involves the direct transition from one C site to another, with an activation energy of 0.81 eV. In α-Fe, the dimer method has been applied to study the directional change of nanosize interstitial clusters, which is inaccessible to classical MD. The activation energies and mechanisms for the directional change strongly depend on the cluster size. The directional change for small clusters ($N = 1, 2,$ and 3) can be thermally activated at room temperature, but clusters of size 4 and 5 may change their direction only at high temperatures. The energy barriers for directional change of a cluster size

larger than 5 interstitials are in excess of 2 eV, so that it is very difficult for it to be activated thermally. Small clusters change their direction via <110> dumbbell mechanism involving rotations into and out of the <110> dumbbell configurations, whereas the directional change for larger clusters is a two-step process consisting of a translation along the <100> direction and rotation into a equivalent <111> configuration.

Based on the dimer method, long time dynamics has been developed [78]. As soon as the saddle point searches are carried out, the rate for each transition mechanism is estimated using harmonic transition state theory. Transitions are selected and the clock is advanced according to the kinetic Monte Carlo algorithm. Unlike traditional applications of kinetic Monte Carlo, the atoms are not assumed to sit on lattice sites and a list of all possible transitions need not be specified beforehand. It is possible to employ this method to study the formation of nanostructures induced by ion–solid interaction.

6 Summary and Outlook

State-of-the-art computational methods to simulate defect configurations, nanoscale structures and the microstructural changes under ion implantation or irradiation are reviewed. It has been demonstrated that the evolution of defects and nanoscale structures is a multiscale phenomenon, which involves processes spanning a wide range of length and time scales. As such, different computational methods with different spatial and time scales have been employed to understand this multiscale phenomenon, which includes ab initio electronic structure calculations, molecular dynamics (MD) simulations, long-time scale dynamics, and kinetic Monte-Carlo (KMC) methods. Ab initio calculations generally require large computational efforts compared to the MD and KMC simulations, but provide a set of accurate defect properties for the development of interatomic potentials that are important for the upscaling atomic-level simulations.

Although these methods have been successfully applied to understand fundamental aspects of defect properties, defect accumulation, the formation of nanoscale structures, defect annealing, and recovery, the realism and accuracy of empirical potentials in MD simulations, developed by fitting to lattice constants and elastic constants at equilibrium, is limited, electronic effects are not well-represented, and charge transfer and charge-density redistribution are ignored. With increasing computer power, ab initio molecular dynamics (AIMD) simulations of defect production have recently emerged to address these scientific issues and understand the coupling of electronic and atomic dynamics. To demonstrate the feasibility of such simulations, ion–solid interactions at low energies have been investigated using AIMD in GaN and SiC [79, 80], and the calculated values of displacement threshold energies in SiC and GaN differ to varying extents from those given by empirical potential calculations. The AIMD simulations may provide an explanation for observed phenomena such as nitrogen bubble formation in GaN under electron irradiation and

may also shed light on the influence of charge transfer on defect dynamics. A major computational challenge to be undertaken is to directly simulate defect generation and formation of nanoscale structures due to ion–solid interaction and defect kinetics using the AIMD method. These studies should provide more accurate information for upscaling simulations. As computational speed and code performance improve, it is expected that higher-energy displacement cascades can be directly simulated using the AIMD method within the next ten years, which will significantly improve our fundamental understanding of electronic and atomic processes on defect dynamics, phase equilibrium, and nano/microstructure evolution.

References

1. T. Diaz de la Rubia, R. S. Averback, R. Benedek, W. E. King, Phys. Rev. Lett. 59 (1987) 1930.
2. D. J. Bacon, T. Diaz de la Rubia, J. Nucl. Mater. 216 (1994) 275.
3. D. J. Bacon, F. Gao, Yu. N. Osetsky, J. Nucl. Mater. 276 (2000) 1.
4. J. M. Perlado, L. Malerba, A. Sánchez Rubio, T. Díaz la Rubia, J. Nucl. Mater. 276 (2000) 235.
5. F. Gao, W. J. Weber, R. Devanathan, Nucl. Instrum. Methods B 180 (2001) 176.
6. K. Nordlund, Nucl. Instrum. Methods B 188 (2002) 41.
7. C. Meis, J. D. Gale, Mater. Sci. Eng. B 57 (1998) 52.
8. B. Park, W. J. Weber, L. René Corrales, Phys. Rev. B 64 (2001) 174108.
9. E. Wendler, A. Helt, W. Wesch, Nucl. Instrum. Methods B 141 (1998) 105.
10. W. J. Weber, N. Yu, L. M. Wang, N. J. Hess, Mater. Sci. Eng. A 253 (1998) 62.
11. P. Hohenberg, W. Kohn, Phys. Rev. 136 (1994) B864.
12. W. Kohn, L. J. Sham, Phys. Rev. 137 (1965) A1697.
13. M. D. Segall, P. J. D. Lindan, M. J. Probert, C. J. Pickard, P. J. Hasnip, S. J. Clark, M. C. Payne, J. Phys. Condens. Matter 14 (2002) 2717.
14. M. D. Segall, J. Phys. Condens. Matter 14 (2002) 2957.
15. J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, D. Sánchez-Portal, J. Phys. Condens. Matter 14 (2002) 2745.
16. D. R. Bowler, T. Miyazaki, M. J. Gillan, Comput. Phys. Commun. 137 (2001) 255.
17. B. J. Coomer, J. P. Goss, R. Jones, S. Öberg, P. R. Briddon, J. Phys. Condens. Matter 13 (2001) L1.
18. A. F. Wright, Phys. Rev. B 74 (2006) 165116.
19. A. Mattausch, M. Bockstedte, O. Pankratov, Mater. Sci. Forum 353–356 (2001) 323.
20. F. Gao, E. J. Bylaska, W. J. Weber, L. R. Corrales, Phys. Rev. B 64 (2001) 125203.
21. J. E. Northrup, S. B. Zhang, Phys. Rev. B 47 (1993) 6791.
22. S. Pöykkö, M. J. Puska, R. M. Nieminen, Phys. Rev. B 53 (1996) 3813.
23. J. Neugebauer, C. G. Van de Walle, Phys. Rev. B 50 (1994) 8067.
24. F. Gao, E. J. Bylaska, W. J. Weber, Phys. Rev. B 70 (2004) 125203.
25. A. Chartier, C. Meis, W. J. Weber, L. René. Corrales, Phys. Rev. B 65 (2002) 134116.
26. H.Y. Xiao, L. M. Wang, X. T. Zu, J. Lian, R. C. Ewing, J. Phys. Condens. Matter 19 (2007) 346203.
27. I. Gorczyca, A. Svane, N. E. Christensen, Phys. Rev. B 60 (1999) 8147.
28. P. Boguslawski, E. L. Briggs, J. Bernholc, Phys. Rev. B 51 (1995) 17255.
29. T. Mattila, R.M. Nieminen, Phys. Rev. B 55 (1997) 9571.
30. F. Gao, E. J. Bylaska, A. El-Azab, W. J. Weber, Appl. Phys. Lett. 85 (2004) 5565.
31. W. Windl, T. Lenosky, J. Kress, A. Voter, Nucl. Instrum. Methods B 141 (1998) 61.
32. G. Lucas, L. Pizzagalli, Phys. Rev. B 72 (2005) 161202.
33. F. Ercolessi, J. B. Adams, Europhys. Lett. 26 (1994) 583.

34. A. M. Raphuthi, X. Q. Wang, F. Ercolessi, J. B. Adams, Phys. Rev. B 52 (1995) R5554.
35. J. D. Rittner, D. N. Seidman, K. L. Merkle, Phys. Rev. B 53 (1996) R4241.
36. D. Y. Sun, X. G. Gong, Phys. Rev. B 57 (1998) 4730.
37. F. Gao, W. J. Weber, Nucl. Instrum. Methods B 191 (2002) 504.
38. M. Posselt, F. Gao, W. J. Weber, V. Belko, J. Phys. Condens. Matter 16 (2004) 1307.
39. J. B. Gibson, A. N. Goland, M. Milgram, G. H. Vineyard, Phys. Rev. 120 (1960) 1229.
40. J. R. Beeler, Jr., D. G. Besco, Radiation Damage in Solids, Vol. I, I.A.E.A., Vienna, 1962, p. 105.
41. T. Diaz de la Rubia, R. S. Averback, R. Benedek, W. E. King, Phys. Rev. Lett. 59 (1987) 1930, see also erratum: Phys. Rev. Lett. 60 (1988) 76.
42. M. M. Jakas, D. E. Harrison, Jr., Phys. Rev. Lett. 55 (1985) 1782.
43. R. Smith, Proc. Roy. Soc. Lond. A 431 (1990) 143.
44. H. M. Urbassek, K. T. Waldeer, Phys. Rev. Lett. 67 (1991) 105.
45. R. Smith, R. P. Webb, Nucl. Instrum. Methods B 59/60 (1991) 1378.
46. R. W. Hockney, J. W. Eastwood, Computer Simulations Using Particles, Adam Hilger, Bristol, 1988.
47. W. C. Swope, H. C. Andersen, P. H. Berens, K. R. Wilson, J. Chem. Phys. 76 (1982) 637.
48. C. W. Gear, Numerical Initial Value Problems in Ordinary Differential Equations, Prentice-Hall, Englewood Cliffs, NJ, 1971.
49. R. Smith, Atomic and Ion Collisions in Solids and at Surfaces, Cambridge University Press, Cambridge, UK, 2005.
50. D. J. Bacon, A. F. Calder, F. Gao, V. G. Kapinos, S. J. Wooding, Nucl. Instrum. Methods B. 102 (1995) 37.
51. D. J. Bacon, F. Gao, Yu. N. Osetsky, J. Nucl. Mater. 276 (2000) 1.
52. F. Gao, D. J. Bacon, L. M. Howe, C. B. So, J. Nucl. Mater. 294 (2001) 288.
53. F. Gao, W. J. Weber, J. Appl. Phys. 89, (2001) 4275.
54. R. Devanathan, W. J. Weber, T. Diaz de la Rubia, Nucl. Instrum. Methods B 141 (1998) 118.
55. R. Kilaas, NECM HRTEM Image Simulation Software, National Center for Electron Microscopy, Lawrence Berkeley Laboratory, Berkley, CA.
56. Y. Zhang, W. J. Weber, W. Jiang, A. Hallen, G. Possnert, J. Appl. Phys. 91 (2002) 6388.
57. W. J. Weber, Nucl. Instrum. Methods B 166–167 (2000) 98.
58. W. Jiang, W. J. Weber, S. Thevuthasan, V. Shutthanandan, J. Nucl. Mater. 289 (2001) 96.
59. W. Jiang, W. J. Weber, S. Thevuthasan, D. E. McCready, Nucl. Instrum. Methods B 143 (1998) 333.
60. F. Gao, W. J. Weber, Phys. Rev. B 66 (2002) 024106.
61. W. J. Weber, N. Yu, L. M. Wang, J. Nucl. Mater. 253 (1998) 53.
62. W. J. Weber, L. M. Wang, Nucl. Instrum. Methods B 106 (1995) 298.
63. H. L. Heinisch, B. N. Singh, J. Nucl. Mater. 251 (1997) 77.
64. F. Gao, D. J. Bacon, A. V. Barashev, H. L. Heinisch, Mater. Res. Soc. Symp. Proc. 540 (1999) 703.
65. H. L. Heinisch, B. N. Singh, J. Nucl. Mater. 271–272 (1999) 46.
66. M. J. Caturla, N. Soneda, E. Alonso, B. D. Wirth, T. Díaz de la Rubia, J. Nucl. Mater. 276 (2000) 13.
67. L. Pelaz, L. A. Marques, M. Aboy, G. Gilmer, L. A. Bailon, J. Barbolla, Comput. Mater. Sci. 27 (2003) 1.
68. H. L. Heinisch, B. N. Singh, Phil. Mag. 83 (2003) 3661.
69. J. H. Evans, Phil. Mag. 86 (2006) 173.
70. G. Otto, G. Hobler, Mater. Res. Soc. Symp. Proc. 792 (2004) 525.
71. Z. Rong, F. Gao, W. J. Weber, G. Hobler, J. Appl. Phys. 102 (2007) 103508.
72. W. Jiang, W. J. Weber, S. Thevuthasan, D. E. McCready, Surf. Interface Anal. 27 (1999) 179.
73. Y. Zhang, W. J. Weber, W. Jiang, C. M. Wang, V. Shutthanandan, A. Hallen, J. Appl. Phys. 95 (2004) 4012.

74. M. R. Sørensen, A. F. Voter, J. Chem. Phys. 112 (2000) 9599.
75. B. P. Uberuaga, R. Smith, A. R. Cleave, F. Montalenti, G. Henkelman, R. W. Grimes, A. F. Voter, K. E. Sickafus, Phys. Rev. Lett. 92 (2004) 115505.
76. G. Henkelman, H. Jónsson, J. Chem. Phys. 111 (1999) 7010.
77. F. Gao, G. Henkelman, W. J. Weber, L. R. Corrales, H. Jónsson, Nucl. Inst. Methods B 202 (2003) 1.
78. G. Henkelman, H. Jónsson, J. Chem. Phys. 115 (2001) 9657.
79. H. Y. Xiao, F. Gao, X. T. Zu, W. J. Weber, Ab initio molecular dynamics study of the threshold displacement energy in GaN, J. Appl. Phys. 105 (2009) 123527.
80. F. Gao, H. Y. Xiao, X. T. Zu, M. Posselt, W. J. Weber, Defect-enhanced charge transfer by ion-solid interaction in SiC using large-scale ab initio molecular dynamics simulations, Phys. Rev. Lett. 103 (2009) 027405.

Characterizing Nanoscale Crystal Perfection by Crystal Mapping

Sachiko T. Nakagawa

1 Introduction

Crystal defect engineering is driven by the need to minimize defects with negative consequences, e.g., in photo-electronic devices [1] or in promoting the aggregation of defects to create nanoparticles [2].

In order to develop an understanding of the different aspects of defect engineering, research is directed towards understanding the entire chain of interactions between the projectile and the material, from the collision to the relaxation stage.

The whole process can be described well by computer simulations, as will be discussed in chapter 8, "Computer Simulation Methods for Defect Configurations and Nanoscale Structures." Here, the characterization of the resultant *lattice defects* is discussed. The different types of lattice defects are important because: point defects can modify the ligand geometry leading to, e.g., deactivation of a dopant; linear and planar defects can disturb nanometer-scale fabrication of ultra-fine devices; the nanoclusters are important for photo-electronic devices; cluster formation may be sensitive to thermally persistent defects; amorphization or structure change can influence all these processes as a whole; and so on. Defects in amorphous material, such as the columnar defect [3], are not considered in this chapter.

Part of the kinetic energy introduced by the projectile is released outwards, by projectile reflection, sputtering, secondary electron emission, photon emission, or thermal energy flow through the heat-bath. The residual deposited energy will eventually be transferred to the lattice (or cohesive) system, causing defect formation [4]. This can take place by direct or indirect processes when electron-phonon interactions intervene [5, 6]. Generally speaking, the longer the elapsed time of energy dissipation in the collision process, the wider the region expands, with a more diluted energy deposition.

For computational physics investigations of defect formation, Monte Carlo (MC) or molecular dynamics (MD) simulations are often used within the framework

S.T. Nakagawa (✉)
Graduate School of Science, Okayama University of Science, Okayama 700-0005, Japan
e-mail: stnak@dap.ous.ac.jp

of classical dynamics. MC is most useful to investigate the knock-on processes that form point defects. For higher order defect formation, however, energy transfers below the displacement threshold are very important. Extended defects are formed in the relaxation stage where MD methods must be used, because MD takes into account multi-collision processes during the time evolution of relaxation. Hybridized simulations that combine MC and MD are useful in this context because MD calculations are time-consuming, especially when high-energy knock-on processes need to be modeled [7].

In order to comprehend the defect formation process, dynamic simulation and analysis of the defects produced should be carried out as concurrently as possible. It is possible to visualize the defects produced in an MD super-cell. However, such two-dimensional projections and time sequence visualizations give little or no crucial predictive information. In order to obtain information about lattice defects, where they are located, how they correlate to each other, how the ligand surrounding them is related, and how they relate to the host material, a crystallographic description of the defects is indispensable.

In defect analysis, two characteristics are particularly worthy of consideration: one is the "range-order" and the other is "individuality of defects." The *range-order* refers to the distorted size. There are three ranges [8, 9], i.e., the long-range-order (LRO), medium-range-order (MRO) [10], and short-range-order (SRO), as are shown in Table 1. SRO usually includes the first and second nearest neighboring atoms, MRO spans a region up to 2–3 nm at most [11], and LRO involves the entire space being studied. Note that LRO is established by the sequence of SRO. On the other hand, the "individuality of defects" relates to how the defects look independently. Some defects should be considered individually, whereas other defects are significant only after amorphization or a structural change. Therefore, we arrange defects according to these two categories in this chapter.

Table 1 Lattice defects to be resolved by computation

		SRO	LRO	
			MRO	LRO
Section 2.3: No-integration (individual)		Point defects (PM, LS, AWS)	Planar defect (PM) Linear defect (PM)	Domain-structure (PM) Interface (PM)
Section 2.4: Integration (averaged)	With unit	Deformations (Voronoi-analysis, Ring-statistics)	Amorphization and structure change (Voronoi-analysis, Ring-statistics)	
	Without-unit	Deformations ($g(r)$) Minimum yield χ_{min} (PM)	Amorphization (PM, BW, $g(r)$), Structure changes (PM, $g(r)$) Relative intensity of RBS-C signal (PM)	

The abbreviations are PM (Pixel mapping), LS (Lindeman sphere), AWS (Atomic-Wigner-Seitz), BW (Bragg-Williams), and $g(r)$ (radial distribution function). See the text in Sects. 2.3 and 2.4

2 Methods of Lattice-Defect Characterization

2.1 Conventional Description of Atoms in a Crystal

Any lattice defect can be considered to be a deviation from the ideal configuration of atoms in a perfect crystal. There are many polymorphic crystals with the same stoichiometry, as will be seen in Table 2 later. In order to consider structural changes from one crystal configuration to another, as many as possible candidate prototype crystals should be considered in analysis.

A crystal shows translational symmetry established by repetition of Bravais lattices. The nature of a crystal is determined by the atomic basis located at each Bravais lattice point. Mathematically this situation is expressed as the sum of the primitive vector $\{P_j\}$ ($j = 1$–3) and the basis vector $\{B_k\}$ ($k = 1 - N_m$) [12, 13]. The *space lattice* is spanned by the set of $\{P_j\}$ that defines *primitive lattice points*. A single atomic basis is described by the set of vectors $\{B_k\}$, which, as noted above, is repeated at each primitive lattice point in space to form a perfect crystal. That is, each atom and each lattice site maintain a one-to-one correspondence. Each component of $\{B_k\}$ is usually expanded by the set of $\{P_j\}$. Note that in general components of $\{P_j\}$ and $\{B_k\}$ are usually neither normalized nor orthogonal to each other.

The description above is complete for a perfect crystal excluding thermal vibrations. However, it is insufficient to extract any crystallographic information from an ensemble of atoms in an actual simulation. It follows that crystallographic identification, such as for linear or planar defects, is almost impossible. The two key issues that need to be addressed to achieve identification are: (1) In reality every atom should have a marginal volume to be recognized. This margin corresponds to the resolution power of the instruments and the statistical smearing in computation. (2) It is impossible to uniquely decompose a position vector in the six-dimensional space made of $\{P_j\}$ and $\{B_k\}$.[1] The Pixel Mapping (PM) method [14, 15] that is explained in the next section has solved these two problems, to the extent that one can even exclusively identify the Miller index of planar defects or their crystallographic direction in terms of an atomic row.

2.2 Crystallographic Analysis: Pixel Mapping for Cubic Crystals

In crystallography, inorganic crystals in nature are classified with *sequential international numbers* from 1 through 230 defining the space groups [16]. The space group is mathematically complete; however, not every space group has a one-to-one correspondence with a crystal in nature. Inorganic crystals are often alternatively

---

[1] Moreover, it is arbitrary where the origin of $\{B_k\}$ is placed in one unit cell defined by $\{P_i\}$. For example, in some cases, which are indicated by "**" in the last second column in Table 2, for convenience it is shifted from the SB [13] definition.

Table 2 The 24 prototype crystals in which PM can be used [15]

No.	Space group(#)	Primitive vectors	Prototype	Type[a]	Basis vectors (p_k, q_k, r_k) in Eq. (1) defining the geometry of the "*atomic-basis*", with $n_j = 4$, except for the prototypes No. 19, 20 ($n_j = 8$)		Type name
1	216	fcc[b]	ZnS	AB	A (0,0,0), B (1,1,1)		Zincblende
2			MgAsAg	ABC	A (0,0,0), B (2,0,0), C (1,1,1)		Half-Heusler
3	221	sc[c]	αPo	A	A (0,0,0)		sc
4			CsCl	AB	A (0,0,0), B (2,2,2)		
5			ReO$_3$	AB$_3$	A (0,0,0), B (2,0,0)[d]		αReO$_3$
6			CaTiO$_3$	ABC$_3$	A (0,0,0), B (2,2,2), C (0,2,2)[d]		Perovskite
7			Cu$_3$Au	AB$_3$	A (0,0,0), B (0,2,2)[d]		
8			NbO	AB	A (0,2,2)[d], B (0,0,2)[d]		
9	223		Cr$_3$Si	AB$_3$	A (0,0,0), A (2,2,2), even B (0,1,2)[d], odd B (0,2,3)[d]		Cr$_3$Si
10	224		Ag$_2$O	A$_2$B	A (0,0,0), A (0,2,2)[d], B (1,1,1), B (3,3,3)		Cuprite
11	225	fcc[b]	Cu	A	A (0,0,0)		fcc
12			NaCl	AB	A (0,0,0), B (2,2,2)		Rock-salt
13			CaF$_2$	AB$_2$	A (0,0,0), B (1,1,1), B (3,3,3)		Fluorite
14			AlFe$_3$	AB$_3$	A (0,0,0), B (1,1,1), B (2,0,0), B (3,3,3)		AlFe$_3$
15			AlCu$_2$Mn	AB$_2$C	A (0,0,0), B (1,1,1), B (3,3,3), C (2,0,0)		Heusler
16			Ca$_7$Ge	AB$_7$	A (0,0,0), B (0,1,1)[d], even B (013)[d], B (2,2,2)		
17	227		C	A	A (0,0,0), A (1,1,1)	–[e]	Diamond
18			NaTl	AB	A (0,0,0), A (1,1,1), B (2,2,2), B (3,3,3)	–[e]	
19		fcc[f]	SiO$_2$	AB$_2$	A (0,0,0), A (2,2,2), B (1,1,1), B (1,3,3)[d] ($n_j = 8$)	–[e]	β-cristobalite
20			Cu$_2$Mg	AB$_2$	A (0,0,0), A (2,2,2), B (3,3,5)[d], B (5,5,5) ($n_j = 8$)	–[e]	Laves
21	229	bcc[g]	W	A	A (0,0,0)		bcc
22			Pt$_3$O$_4$	A$_3$B$_4$	A (0,2,2)[d], B (1,1,1), B (1,1,3)[d]		
23			La$_2$O$_3$	A$_2$B$_3$	A (0,0,0), B stays at (200)[d] with 50% probability		
24			CrFe$_4$Ni$_3$	AB$_4$C$_3$	A (0,0,0), B (1,1,1), B (1,1,3)[d], C (0,0,2)[d]		Hypo. Ferrite

[a] *Indicating the elemental composition formula of a crystal, whose prototype is shown in the rightmost column.*

[b] P_1 (0, 2, 2), P_2 (2, 0, 2), P_3 (2, 2, 0) with $n_j = 4$.

[c] P_1 (4, 0, 0), P_2 (0, 4, 0), P_3 (0, 0, 4), with $n_j = 4$.

[d] Expressing (all) possible permutations.

[e] Indicating different expressions from SB [15].

[f] P_1 (0, 4, 4), P_2 (4, 0, 4), P_3 (4, 4, 0), with $n_j = 8$.

[g] P_1 (2, 2, 2), P_2 (2, -2, 2), P_3 (2, 2, -2) with $n_j = 4$.

classified by representative *prototype*s in a database called *StrukturBericht* (SB) [13]. A prototype crystal is defined by crystal structure, the ligand field, and stoichiometry; thus it follows that in the same prototype group, crystals have a very similar physicochemical nature.

From the framework of the SB database [13], we have observed that crystals can be rearranged by a regular segmentation rule, if irregular cases are defined [15]. The regular case means that $\{B_k\}$ is expressed by a set of $\{P_j\}$ using coefficients of rational numbers. So far we have concentrated on cubic crystals[2] that belong to the space group from 195 to 230. Therein a rule was defined to segment the Bravais lattice into $4 \times 4 \times 4$ for simple crystals and $8 \times 8 \times 8$ for superimposed crystals. We have called this PM segmentation. From the nature of the $\{P_j\}$ vectors, the 24 regular crystals listed in Table 2[3] were classified into three groups: Table 3a is for "*fcc*-family crystals" with ($n_j = 4$), Table 3b for "*bcc*- and *sc*-family crystals" with ($n_j = 4$), and Table 3c for "special *fcc*-family crystals" with ($n_j = 8$). The sheets making up Tables 3a, 3b, and 3c we call the PM table hereafter.

The segmentation changes the unit length to describe a crystal, from the "lattice constant d" to another unit $l_p = d/n_j$, where $n_j = 4$ or 8. Hence, both B_k and P_j are described by orthogonal unit-vectors $\{e_x, e_y, e_z\}$, as follows

$$B_k = p_k \cdot (l_p \cdot e_x) + q_k \cdot (l_p \cdot e_y) + r_k \cdot (l_p \cdot e_z)$$
$$= p_k \cdot (d / n_j)e_x + q_k \cdot (d / n_j)e_y + r_k \cdot (d / n_j)e_z, \tag{1}$$
$$\text{or simply, } B_k = (p_k, q_k, r_k).$$
$$P_j = p_j \cdot (l_p \cdot e_x) + q_j \cdot (l_p \cdot e_y) + r_j \cdot (l_p \cdot e_z)$$
$$= p_j \cdot (d / n_j)e_x + q_j \cdot (d / n_j)e_y + r_j \cdot (d / n_j)e_z,$$
$$\text{or simply, } P_j = (p_j, q_j, r_j). \tag{2}$$

All components of both the primitive vector $P_j = (p_j, q_j, r_j)$ and the basis vector $B_k = (p_k, q_k, r_k)$ are integer numbers for all *regular crystals* listed in Table 2. For example, components of $\{B_1, B_2, , B_{Nm}\}$ are shown in Table 2.[4] The number N_m is $N_m = 2$ for A of a monatomic crystal (No. 17), $N_m = 6$ for AB of a binary crystal (No. 8), $N_m = 8$ for AB_3 of a ternary crystal (No. 9), or $N_m = 2.5$ for A_2B_3 (No. 23), and so on.

Consequently, in PM, the location of an atom at a point $X(x, y, z)$ is expressed in terms of an integer address (I_x, I_y, I_z) of a pixel that accommodates the atom within the finite volume of $l_p \times l_p \times l_p$. With the Gaussian notation "[ ]," the location is expressed as follows:

---

[2] In cubic crystals, the crystal plane ($\eta\kappa\lambda$) and crystal direction [$\eta\kappa\lambda$] are perpendicular to each other. This property helps us to straightforwardly identify planar and linear defects [17], within the same vector analysis framework.

[3] The "prototype" selection offers considerable advantages to deriving a new interatomic potential from established formulas for atoms in the same prototype group [15].

[4] In the 24 prototype crystals listed in Table 2, the order of atoms *A, B,* and *C* in $A_l B_m C_n$ is chosen so that a type *A* atom is located at the corner of a Bravais lattice as often as possible.

Table 3a Atomic configuration in each prototype crystals with relevant LRO components [15]. A set of fractions in each row for relevant prototype shows the LRO components, while fractions in the row of LRO (∞) represent each component of LRO in the amorphous extreme. *fcc*-family crystals with ($n_j = 4$)

Remainders (pixel group)		Zero $l+m+n=4k'$		One $4k'+1$		Two $4k'+2$		Three $4k'+3$	
Label and parity		(1): all even	<1'> others	{T}: all odd	<T'> others	{H}: all even	<H'> others	(2): all odd	<2'> others
Sites (l, m, n)		(0, 0, 0), (0, 2, 2)[a]	Others	(1,1,3)[a], (3,3,3)[a]	Others	(0,0,2)[a], (2,2,2)	Others	(1,1,1), (1,3,3)[a]	Others
LRO(∞) parameter		4/64	12/64	4/64	12/64	4/64	12/64	4/64	12/64
No.	Type[b] Prototype								
1	AB ZnS	A: 1/2						B: 1/2	
2	ABC MgAsAg	A: 1/3				C: 1/3		B: 1/3	
11	A Cu	A: 1/1							
12	AB NaCl	A: 1/2				B: 1/2			
13	AB$_2$ CaF$_2$	A: 1/3		B: 1/3				B: 1/3	
14	AB$_3$ AlFe$_3$	A: 1/4		B: 1/4		B: 1/4		B: 1/4	
15	AB$_2$C AlCu$_2$Mn	A: 1/4		B: 1/4		C: 1/4		B: 1/4	
16	AB$_7$ Ca$_7$Ge	A: 1/8	B: 3/8			B: 1/8	B: 3/8		
17	A C	A: 1/2						A: 1/2	
18	AB NaTl	A: 1/4		B: 1/4		B: 1/4		A: 1/4	

[a] Expressing (all) possible permutation.
[b] Indicating the elemental composition formula of a crystal, whose prototype is shown in the left column.

Table 3b Atomic configuration in each prototype crystals with relevant LRO components [15]. A set of fractions in each row for relevant prototype shows the LRO components, while fractions in the row of LRO (∞) represent each component of LRO in the amorphous extreme. bcc- and sc-family crystals with ($n_j = 4$)

Remainders (pixel group)	Zero $l+m+n=4k'$			One $4k'+1$			Two $4k'+2$			Three $4k'+3$		
Label and parity	(1): all even		$<1>$ others	{T}: all odd		$<T'>$ others	{H}: all even		$<H'>$ others	(2): all odd		$<2'>$ others
	$(1)_1$	$(1)_3$		${T}_1$	${T}_3$		${H}_1$	${H}_3$		$(2)_1$	$(2)_3$	
Sites (l,m,n)	$(0,0,0)$	$(0,2,2)^a$	Others	$(3,3,3)$	$(1,1,3)^a$	Others	$(2,2,2)$	$(0,0,2)^a$	Others	$(1,1,1)$	$(1,3,3)^a$	Others
LRO(∞) parameter	1/64	3/64	12/64	1/64	3/64	12/64	1/64	3/64	12/64	1/64	3/64	12/64

No.	Type[b]	Prototype	$(1)_1$	$(1)_3$	$<1>$	${T}_1$	${T}_3$	$<T'>$	${H}_1$	${H}_3$	$<H'>$	$(2)_1$	$(2)_3$	$<2'>$
3	A	αPo	A: 1/1											
4	AB	CsCl	A: 1/2						B: 1/2					
5	AB$_3$	ReO$_3$(α)	A: 1/4							B: 3/4				
6	ABC$_3$	CaTiO$_3$	A: 1/5	C: 3/5					B: 1/5					
7	AB$_3$	Cu$_3$Au	A: 1/4	B: 3/4										
8	AB	NbO		A: 1/2					B: 1/2					
9	AB$_3$	Cr$_3$Si	A: 1/8					B odd $(023)^a$: 3/8	A: 1/8					B-even $(012)^a$: 3/8
10	A$_2$B	Ag$_2$O	A: 1/3			B: 1/3						B: 1/3		
21	A	W	A: 1/2						A 1/2					
22	A$_3$B$_4$	Pt$_3$O$_4$		A: 3/14		B: 4/14				A: 3/14		B: 4/14		
23	A$_2$B$_3$	La$_2$O$_3$	A: 2/10 B(50%): 3/10						A: 2/10 B(50%): 3/10					
24	AB$_4$C$_3$	CrFe$_4$Ni$_3$	A: 1/16 C: 3/16			B: 1/16	B: 3/16		A: 1/16 C: 3/16			B: 1/16	B: 3/16	

a Expressing (all) possible permutation.
b Indicating the elemental composition formula of a crystal, whose prototype is shown in the left column.

Table 3c Atomic configuration in each prototype crystals with relevant LRO components [15]. A set of fractions in each row for relevant prototype shows the LRO components, while fractions in the row of LRO (∞) represent each component of LRO in the amorphous extreme. Superimposed-fcc-family crystals with ($n_j = 8$)

Remainders (pixel group)	Zero $l+m+n=8k'$	One $8k'+1$	Two $8k'+2$	Three $8k'+3$	Four $8k'+4$	Five $8k'+5$	Six $8k'+6$	Seven $8k'+7$
Label and parity	(1) all $4k'$ $\Delta=0,4$; $\langle 1\rangle$ others	$\langle 1''\rangle_{64}$ all	$\{T\}$ all even $\Delta=0,4$; $\langle T'\rangle$ others	$\langle T''\rangle_4$ all odd $\Delta=0,4$; $\langle T''\rangle_{12}$ all odd; $\langle T''\rangle_{48}$ others	$\{H\}$ all $4k'$ $\Delta=0,4$; $\langle H'\rangle$ others	$\langle H''\rangle_{64}$ all	(2) all even $\Delta=0,4$; $\langle 2'\rangle$ others	$\langle 2''\rangle_4$ all odd $\Delta=0,4$; $\langle 2''\rangle_{12}$ all odd; $\langle 2''\rangle_{48}$ others
Sites (l,m,n)	(0,0,0), (0,4,4)[a]; Others	All	(6,6,6), (2,2,6)[a]; Others	(1,1,1), (1,5,5)[a]; (1,3,7)[a], (3,3,5)[a], (5,7,7)[a]; Others	(0,0,4)[a], (4,4,4); Others	All	(2,2,2), (2,6,6)[a]; Others	(1,1,5)[a], (5,5,5); (1,3,3)[a], (1,7,7)[a], (3,5,7)[a]; Others
LRO (∞) parameter	4/512; 60/512	64/512	4/512; 60/512	4/512; 12/512; 48/512	4/512; 60/512	64/512	4/512; 60/512	4/512; 12/512; 48/512

No.	Type[b]	Prototype	Zero	Three	Six	Seven
19	AB$_2$	SiO$_2$	A: 1/6	B: 1/6	A: 1/6	B: 1/2
20	AB$_2$	Cu$_2$Mg	A: 1/6	B: 3/6	A: 1/6	B: 1/6

[a] Expressing (all) possible permutations.

[b] Indicating the elemental composition formula of a crystal, whose prototype is shown in the left column.

$$X(x, y, z) \equiv (I_x, I_y, I_z),$$
$$I_x = [x / l_p], I_y = [y / l_p], I_z = [z / l_p], \qquad (3)$$
$$(I_x, I_y, I_z) \equiv T(L, M, N) + u(l, m, n).$$

This transformation means that a "coarse-grained viewpoint" has been introduced. Moreover, it has solved both of the two difficulties that were mentioned in the end of the previous Section (2.1). The volume of $l_p \times l_p \times l_p$ defines a certain marginal area in identifying an atom, which responds to the first difficulty. The decomposition of a position vector $X(x, y, z)$ is unique, as shown in Eq. (3).

2.2.1 PM Table (Tables 3a, 3b, and 3c)

From the set of $u(l, m, n)$ for an atomic ensemble, we can identify its prototype, based on the following *remainder rule*. Each prototype crystal is identified algebraically in terms of the remainder of the sum $(l + m + n)$ divided by n_j and the parity of those integers (l, m, n) as follows. For the pixel labeling, one of the symbols $(1, T, H, 2)$ is used.

The number of assorted pixels for *fcc-family crystals* listed in Table 3a [15] is eight.

$$
\begin{aligned}
l + m + n = 4k' \quad & \text{(1); if } l, m, n \text{ are all even (indicating corner and face-} \\
& \text{centered site),} \\
& <1'>; \text{ otherwise,} \\
l + m + n = 4k' + 1 \quad & \{T\}; \text{ if } l, m, n \text{ are all odd,} \\
& <T'>; \text{ otherwise,} \\
l + m + n = 4k' + 2 \quad & \{H\}; \text{if } l, m, n \text{ are all even; (indicating body-} \\
& and\ side - \text{centered site),} \\
& <H'>; \text{ otherwise,} \\
l + m + n = 4k' + 3 \quad & \text{(2); if } l, m, n \text{ are all odd,} \\
& <2'>; \text{ otherwise.}
\end{aligned}
\qquad (4)
$$

Figure 1 shows the configuration of four kinds of pixels, (1), (2), {T}, and {H}. If (1) and {H} are occupied by Na atoms and Cl atoms, this figure shows a crystal No. 12 (NaCl). If all these four kinds of pixels are fully occupied by a few kinds of atoms, it shows the configuration of a crystal No. 14 or 15 or 18. These four kinds of pixels individually form an *fcc*-lattice. Table 3a describes the perfect allocation of atoms at stable pixels of *fcc*-family crystals in terms of LRO components, as will be explained in Eq. (6) later [15]. A blank space in the PM Table indicates a *non-stable* pixel that is not occupied by an atom in a perfect crystal [15].

The number of assorted pixels for *bcc*- and *sc*-family crystals listed in Table 3b [15] is twelve.

Fig. 1 Configuration of
pixels in unit cell. Four kinds
of pixels of 64 (= 4 × 4 × 4)
pixels are shown in a
conventional unit cell, while
the residual space (75% of
the whole cube) is filled by
unstable pixels [14]. Labels
of pixels are shown (1) by
solid spheres, (2) by *open
spheres*, {T} by *open cubes*,
and {H} by *solid cubes*
Copyright is Nakagawa [14]
as Fig. 2
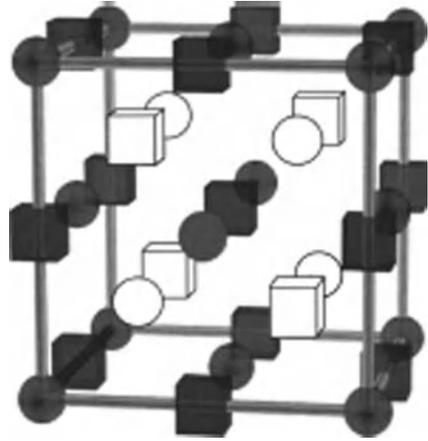
$$l + m + n = 4k' \quad (1)_1; \; l = m = n = 0; \; \text{(indicating corner site)},$$
$$(1)_3; \; \text{if } l, m, n \text{ are all even; (indicating face - centered site)},$$
$$<1'>; \; \text{otherwise},$$
$$l + m + n = 4k' + 1 \quad \{T\}_1; l = m = n = 3$$
$$\{T\}_3; \; \text{if } l, m, n \text{ are all odd},$$
$$<T'>; \; \text{otherwise},$$
$$l + m + n = 4k' + 2 \quad \{H\}_1; \; l = m = n = 2; \; \text{(indicating body centered site)},$$
$$\{H\}_3; \; \text{if } l, m, n \text{ are all even; (indicating side - centered site)},$$
$$<H'>; \; \text{otherwise},$$
$$l + m + n = 4k' + 3 \quad (2)_1; \; l = m = n = 1$$
$$(2)_3; \; \text{if } l, m, n \text{ are all odd},$$
$$<2'>; \; \text{otherwise}.$$

$$(5)$$

Two remarks are added in Table 3b: In crystal No. 9 (AB_3), a B-type atom locates
at either site denoted by odd or even permutation. In crystal No. 23 (A_2B_3), a B-type
atom randomly occupies the half site allowed. Such exceptional assignments are
needed to maintain the stoichiometry of those crystals.

The residual two prototype crystals (Nos. 19 and 20) need a finer segmentation
of $8 \times 8 \times 8 = 512$ because of representing a "superimposed fcc-family." They
are formed by squeezing a smaller size of crystal with half length of the lattice
parameter of the host crystal. The labeling rule can be read from Table 3c. More
detailed information is given in [15]. A further advantage of the PM table is that
it can be used to describe any cubic crystal as a set of parallel atomic planes with
definite Miller indices (*hkl*) and their atomic configurations and to determine the
areal density of atoms on any atomic planes of known area in an MD supercell [18].

2.2.2 Correlation Between PM and MD

PM is conventionally combined with the {NVT} ensembles in MD, i.e., the volume
of MD supercell is invariant. Nevertheless, if the hydrostatic pressure inside the MD

supercell is significant, which can result from surface modification [2] or phase-changes [19], the {NPT} frame should be used instead. In this situation the volume of the MD supercell changes, while the inner pressure and crystallographic structure are fixed. Instead of the trivial redefinition of using "temporary l_p," it is better to use PM at the appropriate stages, when the atomic distribution appears to be stable [15].

Regarding the structure changes, PM can identify the new structures from the change in the LRO components. If structure change occurs locally, PM can be used to define the Miller indices of boundary planes enclosing a specific domain. If a new structure is not identified from PM table, we use a customized approach with real numbers in the components (p_k, q_k, r_k) [15].

2.3 Lattice Defect Analysis: Nonintegrated Information Including Miller Index

We can characterize lattice defects using two categories, as shown in Table 1: One category is the *range-order*, (SRO, MRO, or LRO) [8] and the other is the *individuality of defects* that considers whether the information is obtained as an averaged quantity or not. In this section the individual defects are discussed with four models. They are designated as the PM, LS (Lindeman sphere), AWS (Atomic-Wigner-Seitz cell) [20], and BW (Bragg-Williams) [21, 22] models.

2.3.1 Analysis of Point Defects (Zero-Dimensional)

As a matter of course, individual point defects are rarely experimentally observed except in a few cases: a vacancy (V) can be characterized by electron spin resonance or positron annihilation, and an interstitial (I) trapped at specific sites is resolved from the minimum yield curve in channeling Rutherford Backscattering Spectroscopy (RBS-C) [23].

In computation, point defects are identified by whether the correct atoms are located on the lattice points. Three models (PM, LS, AWS) differently segment cubic crystal space: The PM model segments the space into cubic pixels, where a stable pixel contains one lattice atom at its center. The LS model segments the space into AWS polyhedrons, where a LS locating at the center of AWS accommodates one lattice atom. The segmentation of the AWS model has no specific site for the lattice atom. Both the PM and LS models are more suitable for crystals than the AWS model. Note that the PM and LS models identify an atom from a coarse-grained viewpoint.

As shown in Table 4, the PM and LS approaches similarly identify I and V point defects. In the AWS model, surplus atoms and missing atoms in an AWS are regarded as I and V, respectively. If another kind of constituent atom is trapped at a lattice site, it should be called an antisite (ATS). A foreign atom is an impurity (IMP). The AWS model recognizes another kind of constituent atom as ATS wherever it locates inside the AWS cell, which is different from other models. Hence, it is

Table 4 The dependence of identification of lattice defects upon a model

Model (segmentation)	(i) PM (cube)	(ii) LS (sphere)	(iii) AWS (polyhedron)
V (vacancy)	V in a stable pixel	V in an LS	V in an AWS
I (interstitial)	I in a non-stable pixel	I out of LS	Surplus atom in AWS
Frenkel pair (V + I)	Frenkel pair {(V in a stable pixel) + (I in a non-stable pixel)}	Frenkel pair {V in an LS + I out of LS}	Frenkel pair {(V in an AWS) + (I in another AWS)}
			No defect {(V + I) inside an AWS}
Substitution (IMP)	IMP in a stable pixel	IMP in an LS	IMP in an AWS
Antisite defect (ATS)	WA in a stable pixel	WA in an LS	WA (wherever it locates)
			WA (even V exists)

Different description of point defects by three models. The abbreviations of models are the same as in Table 1. V and I mean vacancy and interstitial, respectively. IMP and WA mean impurity and a wrong atom (another kind of constituent atom), respectively.

an open question whether the number of Frenkel pairs and that of ATS are correctly estimated in the AWS model.

2.3.2 Analysis of Aggregated Defects (1-2-3 Dimensions)

In experiments, extended effects are often aggregated from point defects and are observed with different dimensions [17]. The <110> self-interstitial atom (SIA) chains [24] or crowdion defects peculiar to neutron irradiation [25] are 1-dimensional defects. The antiphase domain structures [10] and {311} planar defects in c-Si [26] are 2-dimensional defects. The nanoparticle [2] and small clusters are 3-dimensional defects. Many theoretical models adopt the Miller index or crystallographic orientation to specify the particular defects. For example, <110> SIA-rods make up of (311) platelets [27], small SIA clusters (<100>- or <110> dimers [28], linear trimer [29], or (100) planar tetramer [30]) for amorphization in c-Si, and ATS clusters for thermally persistent defects in 3C-SiC [31].

In computation, it is easy to identify a 3-dimensional aggregation of point defects. The number of constituent atoms in IMP-clusters [2, 32, 33] or in SIA-clusters [34] is simply counted, assuming a cut-off radius regardless of crystal structure.

However, for 1- or 2-dimensional aggregation, only PM is unique in being able to characterize the crystallographic direction or Miller indices concerned. PM can straightforwardly describe the linear and planar defects using the same framework based on the vector analysis. The Miller index (α, β, γ) of an atomic plane is identified, because integer addresses of atoms composing the plane [15] determine the equation of the plane as $\alpha x + \beta y + \gamma z = D$ with integer D. The normal vector of this plane is $<\alpha, \beta, \gamma>$. The reason for $\Delta D = 1$, including invisible planes in a perfect crystal, is that we expect [14] a migrating interstitial atom to be trapped at

a specific site on a habit plane like a $(311)_{SIA}$ plane. Such a site can be expected to be a "metastable site" that has as high a ligand field as the normal lattice site [18]. The metastable sites are virtually shown behind the Table 2 that tabulates stable sites. Considering the orientation of a dimer, made of two atoms fixed in two pixels (l, m, n) and (l', m', n'), PM simply defines the orientation of the position vector, $(l–l', m–m', n–n')$. In the general case, however, the crystallographic orientation of an atomic row is most simply determined as being parallel to the intersecting vector from the outer products of two normal vectors of the intersecting planes.

2.4 Lattice Defect Analysis: Integrated Information

In this section nonindividual defects are considered with and without the assumption of a minimum unit of a crystal structure. In experiments, the structure change or certain damage in material is detected by means of the broadened signals in diffraction patterns [10], increase of χ_{min} of a minimum-yield curve in curve in RBS-C, increase of relative intensities of RBS-C, or the size change of nanoparticle in, e.g., $L1_0$-FePd [35]. For amorphization and structure change, both the SRO and LRO viewpoints are considered. Both are obtained by numerical integration over the whole region concerned, while it is optional whether some unit cluster is presumed or not.

2.4.1 Analysis with Assuming a Constituent Unit

Each real crystal can be described in terms of an arbitrary unit cluster. They are composed of neighbor atoms, i.e., the so-called nearest-neighbors topology[5] [9], the Voronoi analysis [36, 37], or the Ring-statistics [38, 39]. In such clusters, the modified ligand field is characterized by the coordination number, internuclear distance, and the three-body angles between neighboring atoms.

In a perfect crystal at a temperature of absolute zero K, the Voronoi-polyhedron[6] is identical with an AWS containing an atom at its center. As temperature or damage increases, the shapes of those polyhedrons will be deformed accordingly. The shape is often characterized by the distribution of the indices $(n_3, n_4, \ldots, n_s)$, where n_s indicates the number of polygons with (s) apexes. For example, the Voronoi index is $(0, 12, 0)$ for *fcc* crystal, $(0, 6, 0, 8)$ for *bcc*, and so on. From the change in the distribution of these indices, structural changes can be detected [36, 40].

Another unit of modified ligand fields is the member ring connecting neighbor atoms. For example a perfect c-Si is composed of six-member rings. In damaged c-Si the degree of damage can be qualitatively estimated, from the number distribution of atoms forming rings other than six-member one.

---

[5] The nearest-neighbors topology examines the details of the distorted ligand field surrounding each atom [36].

[6] Voronoi Tessellation in the Delaunay Triangulation (which is often used in the finite element method, if an atom in material is regarded as node point or mother point. Amorphization is identified well by this method [37].

2.4.2 Analysis Without Assuming a Unit

On the other hand, no such unit structure is presumed in other analyses, as arranged in Table 1. By means of the radial distribution function $g(r)$ and LRO parameters defined by the BW (Bragg-Williams) or PM methods, we can characterize structure changes or the degree of crystallinity.

$g(r)$ measures the averaged coordination number as a function of internuclear distance r, irrelevant to the crystallinity. In a perfect crystal $g(r)$ is the superposition of delta functions, which is clearly defined for each prototype crystal using Table 2 and its footnote. For example, in *fcc* crystal, $g(r)$ is 12 and 6 at the first and second nearest neighbor distances, respectively. In *bcc* crystal $g(r) = 8, 6$. If $g(r)$ exhibits changes in its distribution, structural change can be identified, e.g., from *fcc* to *bcc* in an Rb crystal [19]. Usually the asymmetric broadening from delta functions qualitatively reveals the degree of damage in the crystal.

The LRO parameter quantifies the overall degree of randomness in the crystal. Bragg and Williams first defined a single LRO parameter for binary alloy crystals [21, 22]. However, the BW rule cannot be a general definition for all crystals. PM allows unique definition of a set of LRO components for each crystal listed in Table 2. Each component is simply defined as the properly occupied ratio of i-th kind of pixels,

$$N_i = \frac{\text{total number of atoms staying (properly) at } i-\text{th kind of pixels}}{\text{total number of atoms}}, i = 1, j \tag{6}$$

where j is 8, 12, 16 for crystals listed in Tables 3a, 3b, 3c, respectively. The keyword "properly" is used only for stable pixels in chemical compounds, i.e., only the correct atom for the site is counted in Eq. (6) [15]. The set of fractions in each row in the PM table represents LRO components in the perfect state of each prototype crystal as one extreme. The other extreme for an amorphous state is shown as LRO (∞) where all atoms are uniformly distributed. The set of LRO (∞) parameters is common to all prototype crystals in the same sheet, as shown in the first row of each sheet, because any "atomic-basis characteristic of a prototype crystal" looses its identity when atoms spread uniformly. Practically, the degree of crystallinity can be estimated as the intermediate value between those two extremes. Another piece of information that can be read from the LRO parameter relates to the ATS. For monatomic crystals the sum of the LRO components is unity, whatever damage is included. For compound crystals, however, the presence of ATS is indicated when the sum of LRO components is less than unity.

Important information about structure changes can also be obtained from the LRO parameter. If the structure change takes place from one prototype to another within the list in the same sheet of the PM table, the change is easily seen. If structure change occurs between prototypes in different sheets in the PM table, we can identify a new structure after finding the new set of primitive vectors [15].

Because of the coarse-grained viewpoint, PM does not reproduce literally the angular profile of minimum yield obtained from RBS-C. However, PM can roughly estimate the value of χ_{min} in terms of the relevant LRO parameter components. The size of χ_{min} reflects the fraction of IMP trapped at specific interstitial sites located at the center of open channels [23],[7] because the addresses of (l, m, n) of specific pixels that locate at the center of open channels are known. For example, in a crystal No. 17 (diamond type), a <001> axial channel is formed by four atomic rows, the address (l, m, n) of pixels located at the center of <001> channel are "$l + m =$ odd." They are, from the PM table, all nonstable pixels.

3 Characterization of Lattice Defects Using the PM Method

3.1 Analysis Results: Nonintegrated Information Including Miller Index

Point defects (V, I, IMP, ATS, and ATS cluster) were identified making use of PM, using Table 4. PM found a similarity with the LS model, where boron irradiation caused crystalline to amorphous transition in supersaturated c-Si. For example, a sudden decrease of LRO components for stable pixels coincident with a sudden increase of the number of vacancies occurred coincidentally [14]. This result may come from the similarity in the peripheral size of lattice point in both models. On the other hand, we derived a quite different conclusion from the AWS model with respect to ATS for 3C-SiC. The overestimation of ATS led to the wrong conclusion in regard to the trigger of amorphization of a crystal: "ATS is significant on amorphization of 3C-SiC under ion bombardment" [20]. Actually in that case, it turned out interstitials were more important than ATSs [41] for amorphization.

In order to explain the thermally persistent defects produced by ion irradiation in 3C-SiC, a tetrahedral cluster of Si_C-$(C_{Si})_4$ has been proposed from quasistatic calculation [31] based on the smaller formation energy of one type of ATS C_{Si} than Si_C. Nevertheless, the higher formation energy can imply inversely that Si_C can be more stable if it is once formed under energetic ion impacts. In fact, an opposite trend of clear growth of C_{Si}-$(Si_C)_n$ approaching the tetrahedral cluster of C_{Si}-$(Si_C)_4$ appeared where C_{Si}-Si_C bonding expands along $< 111 >$ orientations with 5 keV boron impacts above annealing temperature [42].

Other findings exclusive to the PM method were the determination of the Miller index of atomic planes and rod-like chain structures in them. We reproduced the peculiar aggregation of SIA as a bunch of SIA <011> chains in (311) planes of

---

[7] The randomness to be discussed in χ_{min} differs from that evaluated from relative intensities of RBS-C signals. Because in the dip-curve, many interstitials are hidden, the right atoms are not counted.

c-Si and obtained an areal density of SIA [18] closely similar to that observed in transmission electron microscopy [26].

From the PM table, the so-called double layer structure of {311} planes is found in crystals No. 1 (zincblende) and No. 17 (diamond). Each layer is monatomic and <110> rods aligned in c-Si and paired layer of $\{311\}_C$ and $\{311\}_{Si}$. One period in the sequential alignment of {311} planes is in the order of $(311)_{SIA} - (311)_{Si} - (311)_C - (311)_{SIA}$. The $(311)_{SIA}$ plane is invisible in a perfect crystal; however, migrating SIA would be trapped there [14, 15]. Under boron ion irradiation, we reproduced such a sequence in 3C-SiC [41]. In some cases even $\{311\}_{ATS}$ planes appeared, which is a bunch of well-aligned $<110>_{ATS}$ chains [41].

3.2 Analysis Results: Integrated Information

In PM, without assuming any unit structure, the degree of crystallinity is more quantitatively identified in terms of LRO components (N_i). The correlation between damage accumulation and nanoparticle formation has been confirmed [32]. The evolution in 5 keV B irradiated c-Si was correlated with a definite change in LRO components [14, 34], which finally approached the asymptotic values, LRO(∞). The phase transition was observed as an abrupt change in LRO components in a supersaturated c-Si with boron under a few keV boron impacts at 100 K. At the same time, the number of vacancies coincidently increased. Notably, there was no definite change in $g(r)$, indicating SRO was maintained, at least to some degree [14].

4 Conclusion

A methodology called Pixel Mapping (PM) [14, 15] is demonstrated for characterizing lattice defects atomistically where they are individually characterized and globally, where the distortion-order spans, i.e., short-range, medium-range, or long-range. The 24-prototype cubic crystals were considered. They were taken from the list of natural inorganic crystals in the StrukturBericht [13] database.

An advantage of PM is to present a coarse-grained view of atoms in an MD supercell cube, which makes it possible to determine atomic planes with the Miller index having small indices, and also to determine the crystallographic orientation from numerical data of atomic configurations. Another advantage is to reveal the intermediate steps from one phase (structure) to another in terms of the LRO (long-range-order) components. The PM can characterize point defects including antisites, defect clusters, linear defects with crystallographic orientations, planar defects with Miller indices, and structure changes like amorphization.

References

1. L. Romano, A. M. Piro, M. G. Grimaldi, G. M. Lopez, and V. Fiorentini, Phys. Rev. B **71** (2005) 165201.
2. S. Nakano, H. Ogiso, H. Sato, and S. T. Nakagawa, Surf. Coat. Tech. **128–129** (2000) 71.
3. W. Gerhäuser, G. Ries, H. W. Neumüller, W. Schmidt, O. Eibl, G. Saemann-Ischenko, and S. Klaumünzer, Phys. Rev. Lett. **68** (1992) 879–882.
4. S. T. Nakagawa, *Effect of Disorder and Defects in Ion Implanted Semiconductors, Electrical and Physiochemical Characterization*, Eds. by G. Ghibaudo and C. Christofides, Chap. 3, (Academic Press, Boston, 1997).
5. R. Feynman, *Statistical Mechanics*, Chap. 8, (Westview Press, Boulder, 1972).
6. M. I. Kaganov, I. M. Lifshitz, and L. V. Tanatarov, Sov. Phys. JETP **4** (1957) 173.
7. H. J. Whitlow and S. T. Nakagawa, Nucl. Instrum. Methods B **260/1** (2007) 468.
8. S. R. Ellott, *Physics of Amorphous Materials*, Chaps. 3.3 and 3.4 (Longman Scientific & Technical, Essex, 1990).
9. P. A. Duine, J. Sietsma, and B. J. Thijsse, Phys. Rev. B **50** (1994) 13240.
10. B. E. Warren, *X-ray Diffraction* (Dover Pub., New York, 1990).
11. S. R. Elliott, Nature **354** (1991) 445.
12. C. Kittel,*Introduction to Solid State Physics* (John Wiley & Sons, New York, 1996).
13. Strukturbericht; http://cst-www.nrl.navy.mil/lattice/struk/
14. S. T. Nakagawa, Phys. Rev. B **66** (2002) 094103.
15. S. T. Nakagawa, J. Phys. Soc. Jpn **76** (2007) 034603.
16. A. Schönflies, *Kristallsysteme und Kristallstruktur* (B. G. Teubner, Leipzig, 1891).
17. J. Kim, F. Kirchhoff, J. W. Wilkins, and F. S. Kahn, Phys. Rev. Lett. **84** (2000) 503.
18. S. T. Nakagawa, G. Betz, and H. J. Whitlow, Surf. Coat. Tech. **201/19–20** (2007) 8393.
19. M. Parrinello and A. Rahman, Phys. Rev. Lett. **45** (1980) 1196.
20. F. Gao and W. J. Weber, Phys. Rev. B **66** (2002) 024106.
21. W. L. Bragg and E. J. Williams, Proc. Roy. Soc. (London) A **45** (1934) 699.
22. Z. X. Yang, G. H. Wei, X. Q. Dai, L. D. Wan, M. Wang, and T. Zhang, Phys. Rev. **52** (1995) 10800.
23. E. Yagi, T. Sasahara, T. Joh, M. Hacke, T. Urai, T. Sasamoto, N. Tajime, T. Watanabe, and S. T. Nakagawa, J. Phys. Soc. Jpn. **68** (1999) 4037.
24. S. Takeda, Jpn. J. Appl. Phys. **30** (1991) L639.
25. P. M. Derlet, D. Nguyen-Manh, and S. L. Dudarev, Phys. Rev. B **76** (2007) 054107.
26. S. Takeda and T. Kamino, Phys. Rev. B **51** (1995) 2148.
27. S. Takeda, M. Kohyama, and K. Ibe, Philo. Mag. **70** (1994) 287.
28. G. D. Watkins, R. P. Messmer, C. Weigel, D. Peak, and J. W. Corbett, Phys. Rev. Lett. **27** (1971) 1573.
29. S. K. Estreicher, M. Gharaibeh, P. A. Fedders, and P. Ordejón, Phys. Rev. Lett. **86** (2001) 1247.
30. J. Yamasaki, S. Takeda, and K. Tsuda, Phys. Rev. B **65** (2002) 115213.
31. E. Rauls, Th. Frauenheim, A. Gali, and P. Deák, Phys. Rev. B**68** (2003) 155208.
32. S. T. Nakagawa and G. Betz, Nucl. Instrum. Methods B **180** (2001) 91.
33. S. T. Nakagawa, M. Iwatani, and G. Betz, J. Phys. Soc. Jpn. **75** (2006) 024602.
34. S. T. Nakagawa, Nucl. Instrum. Methods B **206** (2003) 13.
35. K. Sato, Y. Hirotsu, H. Mori, Z. Wang, and T. Hirayama, J. Appl. Phys. **98** (2005) 024308.
36. J. Daligault, Phys. Rev. E **73** (2006) 056407.
37. V. Senthil Kumar and V. Kumaran, Phys. Rev. E **73** (2006) 051305.
38. N. Bernstein and M. J. Aziz, Phys. Rev. B **58** (1998) 4579.
39. K. Kohary, V. M. Burlakov, and D. G. Pettifor, Phys. Rev. B **71** (2005) 235309.
40. M. Tanemura, Y. Hiwatari, H. Matsuda, T. Ogawa, N. Ogita, and A. Ueda, Prog. Theor. Phys. **58** (1977) 1079.
41. S. T. Nakagawa, H. Hashimoto, and G. Betz, Nucl. Instrum. Methods B **258/1** (2007) 194.
42. S. T. Nakagawa, A. Okamoto, and G. Betz, Nucl. Instrum. Methods B **266** (2008) 2711.

Box 4: Interatomic Potential

Fei Gao

In spite of greatly increased computer speeds, the application of ab initio methods for an atomistic simulation of materials is still limited to relatively small ensembles of atoms and, in molecular dynamics, relatively short simulation times ($\sim$5 ps). In contrast, the use of empirical or semiempirical interatomic potentials makes it possible to simulate much larger systems ($\sim$ up to a few million atoms) for much longer times ($\sim$100 ns), and thus to tackle such problems as plastic deformation, ion-solid interaction, or atomic diffusion.

Because of large applications of molecular dynamics methods in materials science, a variety of techniques has been utilized over the years to develop reliable atomic potentials. The earliest classical potentials, mainly employed in the study of liquids, contained only contributions from pair interactions. However, it was soon realized that pair potentials lack certain fundamental properties required for the accurate description of chemical and metallic bonding. One of the early successful attempts to include many-body effects was the introduction of the embedding functional [1], which depends nonlinearly upon the coordination number of each atom. This development leads to the birth of the embedded atom method (EAM) [2, 3], which provides a relatively accurate description of noble transition metals and their alloys. Tersoff potential formulism [4], which is based on the concept of bond order, has been applied to a large number of semiconductors, and Buckingham potential [5] or charge transfer potentials [6] for oxides.

Constructing a potential involves steps (a) and (b) as described below:

(a) Selecting an analytical form for the potential. Novel many-body forms are tried in the attempt to capture as much as possible the physics and chemistry of the bonding. A typical analytical form is constituted by a number of functions, depending on geometrical quantities such as distances or angles, or on intermediate variables such as atom coordinations. For example, a Tersoff potential has the appearance of a pair potential:

F. Gao (✉)
Pacific Northwest National Laboratory, MS K8-87, Richland, WA 99352, USA
e-mail: fei.gao@pnl.gov

$$V = \frac{1}{2} \sum_{ij} \varphi_R \left(r_{ij} \right) + \frac{1}{2} \sum_{ij} B_{ij} \, \phi_A(r_{ij}), \qquad (1)$$

where R and A mean "repulsive" and "attractive." However, it is not a pair potential because B_{ij} is not a constant. In fact, it is the *bond order* for the bond joining i and j, and is a decreasing function of a "coordination" G_{ij} assigned to the bond:

$$B_{ij} = B(G_{ij}) \qquad (2)$$

G_{ij} is in turn defined as

$$G_{ij} = \sum_{k} f_c(r_{ik}) f_c(r_{jk}) g(\theta_{ijk}), \qquad (3)$$

where $f_c(r)$ and $g(\theta_{ijk})$ are suitable functions. The basic idea is that the bond ij is weakened by the presence of other bonds ik involving atom i. The amount of weakening is determined by where these other bonds are placed. Angular terms appear necessary to construct a realistic model.

(b) Finding an actual parameterization for the functions that constitute the analytical form we have chosen. This step is very important and can be technically elaborate.

Variety techniques have been utilized over the years to this end, and fitting data may be obtained from either experimental measurements or first-principles calculations. Commonly, the input data include such quantities for perfect crystal as lattice parameter, cohesive energy, elastic constants, and unrelaxed vacancy formation energy. In all cases, potentials are designed with a "range of applicability" in mind. Due to the vast differences in the electronic structure, it would probably be too ambitious to try modeling a bulk metal and a diatomic molecule of the same element with the same potential: the environment is dramatically different. However, it could be feasible to model simultaneously a bulk and a surface environment, where the environment differs (due to the reduced coordination of the atoms at the surface), but not as dramatically as in the previous example. The ability of a potential to work properly in different environments is called transferability. When using a potential, the simulator should always be familiar with its transferability properties, and validate critically the results obtained in unusual conditions—for example, for very low coordinations, very high temperature, or very high pressure.

A first-principles molecular dynamics program can easily produce huge quantities of useful and reliable data, such as forces acting on atoms in a variety of different geometries and coordinations. These data can complement the experimental quantities normally used to fit potentials, helping to construct more realistic potentials. Recently, a scheme called "force matching method" [7] has been developed. It consists of a numerical optimization procedure which tries to match as closely as possible ab initio-derived forces with a classical potential, using a rich parameterization (10–15 parameters for each function in the potential). By explicitly including

different geometries and different temperatures in the data set, one can attack the transferability problem at its very heart. Based on force matching method, an empirical potential for silicon using the modified embedded atom (MEAM) functional form has been fitted [8]. The model potential has been parameterized using five cubic splines, each with 10 fitting parameters, and a large number of data sets, including ab initio force and energy data on two- to five-atom clusters, liquid and amorphous systems, as well as experimental elastic constants and phonon frequencies. In addition, the energetics for vacancy and interstitial point defects calculated by ab initio LDA (local density approximation) are include in the fit. The potential model provides a reasonable description of energetics for all atomic coordinations, Z, from the dimer ($Z = 1$) to fcc and hcp ($Z = 12$). It accurately reproduces phonons and elastic constants, as well as point defect energetics. It also provides a good description of reconstruction energetics for both the $30°$ and $90°$ partial dislocations. Unlike previous models, this model accurately predicts formation energies and geometries of interstitial complexes—small clusters, interstitial-chain, and planar $\{311\}$ defects.

References

1. J. K. Norskov and N. D. Lang, Phys. Rev. B 21 (1980) 2131.
2. M. S. Daw and M. I. Baskes, Phys. Rev. B 29 (1984) 6443.
3. M. W. Finnis and J. E. Sinclair, Phil. Mag. A 50 (1984) 45.
4. J. Tersoff, Phys. Rev. B 37 (1988) 6991.
5. L. Minervini, R. W. Grimes and K. E. Sickafus, J. Am. Ceram. Soc. 83 (2000) 1873.
6. S. M. Valone and S. R. Atlas, Phil. Mag. 86 (2006) 2683.
7. F. Ercolessi and J. B. Adams, Europhys. Lett. 26 (1994) 583.
8. T. J. Lenosky, B. Sadigh, E. Alonso, V. V. Bulatov, T. Diaz de la Rubia, J. Kim, A. F. Voter and J. D. Kress, Model. Simul. Mater. Sci. Eng. 8 (2000) 825.

Part III
Ion Beam Characterisation of Nanoscale Materials

Ion beams are very sensitive probes for characterising material on a nanometre scale, compared to other probes such as photons and electrons. This is because the ions carry significant momentum and interact through the Coulomb force between the ions and the target electrons and nuclei. This interaction is much stronger than the interaction of the electrons and nuclei with the fields from electomagnetic waves, and the momentum is better preserved than for the scattering of much lighter electrons.

With the development of industrial processes based on nanotechnology, there is an increasing need for materials characterisation for quality control and for process development. For this reason we have restricted the scope of this part to cover only methods, which are able to characterise nanometre structures in an inherently quantitative way. Thus secondary ion mass spectrometry, where the quantitativeness is poor because of the large number of standards that are required, is omitted. Likewise, particle induced X-ray emission (PIXE) is also omitted because, although it is quantitative and sensitive to very small fractions of a monolayer, it is blind to depth information.

Characterising surfaces and thin films is particularly important because materials interact with their environment through the surface layers. It follows that the material's properties are often dominated by the surface properties. For example, the outer layer of metallic atoms such as Rh or Pt is responsible for the action of a catalytic converter. Similarly, understanding the reactions between a metal oxide and silicon is key to tailoring the gate dielectrics for the metal oxide transistors used in the memory and central processors of modern computers and mobile telephones. These methods, which are based on scattering of 100 keV to MeV energy ions, are able to provide compositional information with depth resolution on a scale of 0.1 nm or less. Collective effects, which often are referred as channelling and blocking, provide quite different information about the ordering around defects with a precision that is 0.1 nm or less. The final chapter considers hyperfine interaction methods, which allow measurement of the atomic configuration round defect or impurity atoms with a precision on a sub-Ångström scale. Other interactions between ion beams and the material can also provide important information. Ionoluminescence is particularly interesting because it can also provide an important diagnostic signal during ion beam processing.

Medium Energy Ion Scattering for Near Surface Structure and Depth Profiling

Torgny Gustafsson

1 Introduction and Historical Overview

Medium energy ion scattering (MEIS) is a fairly new and still not very common variant of Rutherford backscattering spectrometry (RBS). The technique is used to give quantitative information about sample composition and surface structure on the nanoscale. Two main quantities can be measured in a MEIS experiment:

- The angular distribution of the backscattered flux from the near-surface region (which gives information about atomic locations (atomic coordinates) in the near-surface region, in favourable cases with 0.001 nm accuracy)
- The energy distribution of the backscattered projectiles (which gives information about the atomic composition of the sample and in particular the depth distribution of the atoms, in favourable cases with ~ 0.1 nm resolution).

Surface crystallography (introduced in Box 5) is still an active area of research with the technique; but today, most work is concentrated on depth profiling of ultrathin films, so far mainly for applications in microelectronics: As ion beams are penetrating, MEIS is one of the very few techniques that can give information about the composition and structure of buried interfaces with monolayer-level resolution.

Compared to conventional RBS, MEIS is performed at lower projectile energies (100–300 keV vs. ~ 2 MeV), and involves more sophisticated/complicated detection equipment. (Sometimes the acronym HRBS (high-resolution RBS) is used, particularly for depth profiling applications at somewhat higher energies than conventional MEIS.) MEIS offers the advantage of superior depth resolution while maintaining the same simple interpretation as RBS. Therefore, an important strength of MEIS is that, as the probe-sample interaction law is known, it is a *quantitative* technique: Scattering yields can be measured in absolute units and compared

T. Gustafsson (✉)

Department of Physics and Astronomy, Rutgers University, Piscataway, NJ 08854, USA
e-mail: gustaf@physics.rutgers.edu

to simulations, also in absolute units. This puts very stringent constraints on the interpretation of the data. Phrased another way, the step between data and interpretation is a very short one, with less room for poorly supported conclusions.

The depth resolution in MEIS decreases with depth, because of ion straggling. Since the scattering cross sections vary as Z^2, the sensitivity to light elements is low. Ion beam methods essentially count the number of atoms in the sample, so there is no chemical specificity. Ion beam-induced damage may however be a problem, but can usually be avoided by rastering the ion beam over the sample. Finally, the detected signal is an average over the beam spot (usually of mm size), so little *lateral* nanoscale information is obtained.

MEIS was first developed by a group at the FOM institute in Amsterdam in the 1970s as a tool for determining surface structure of well characterized, single crystal surfaces in ultra high vacuum. By measuring the angular distribution of the backscattered flux from the surface atoms, specifically the position of so-called "surface blocking dips," and modelling the experiment with Monte Carlo simulations, they were able to determine the position of atoms at simple metal surfaces with an accuracy considerably better than 0.01 nm, fully comparable with results from competing techniques such as Low Energy Electron Diffraction (LEED). In order to isolate the surface signal from scattering from the bulk crystal, very high energy resolution was required. The early work with the technique has been reviewed by van der Veen [1].

The principle of these *channelling and blocking* experiments can be understood from Fig. 1. The incident ion beam is aligned with a high symmetry (channelling) direction of the single crystalline sample. The ion beam can be backscattered from the surface atoms, and, because of vibrations and/or distortions in the near-surface region, also from the first few layers below. Because of the shadowing effects of the surface atoms, the hitting probabilities for atoms away from the surface plane will decrease rapidly. The backscattered flux from the outermost layer will be smooth and featureless. However, the flux from the subsurface layers will show dips
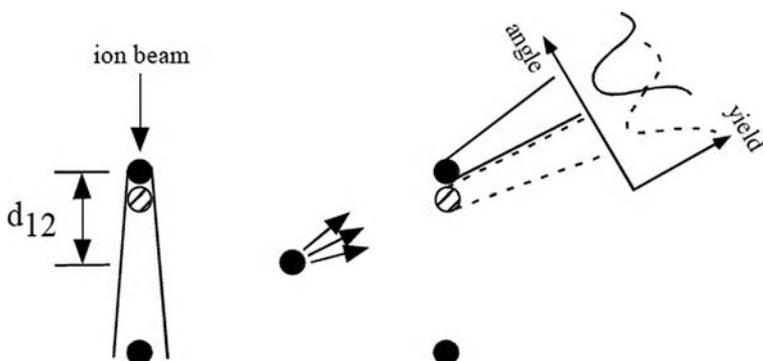
Fig. 1 Schematic illustration of blocking and formation of the blocking dip in the angular spectrum. A change in interlayer spacing is observed experimentally as a shift in the angular position of the blocking dip. The first-to-second interlayer spacing is denoted by d_{12}

(blocking dips) in paths corresponding to the directions from the subsurface to the surface atoms. If the surface layer is displaced outwards (inwards), the surface blocking dip will move towards larger (smaller) scattering angles. Because determining the position of a blocking dip is fairly easy, an accuracy in the determination of the surface layer displacement of 0.01–0.001 nm is possible (see Box 5). However, the ion energy resolution is often (with some notable exceptions, see below) not good enough to allow the separation of the signal from the different surface layers. Blocking dips originating from different layers will therefore overlap. Simple triangulation is in those cases inadequate for an accurate analysis, as discussed further below. However, by merely inspecting the data (say, by looking at the shift of a surface blocking dip relative to its bulk value), it is usually possible to get a good idea of the structural distortions, an important advantage compared to less intuitive methods.

The high energy resolution needed to separate the surface signal from that from the bulk can be used for very high resolution depth profiling. In the mid-1990s, several groups reported results for depth profiling with monolayer or near monolayer resolution. The main systems studied today are oxides deposited on silicon. This was a very timely development: Due to the ever decreasing size of microelectronic components, the semiconductor industry had realized the need to develop an understanding of the silicon oxide/silicon interface on a truly atomistic level (see chapter "High Speed Electronics"). Few techniques were available for quantitative measurements, particularly on buried interfaces. As a replacement for silicon oxide in ever decreasing-sized devices, high dielectric constant oxides (high-k materials) were proposed, but their properties were poorly understood. MEIS was excellently suited for such analysis, and several laboratories quickly reported important results. This stimulated a large expansion of the field. The number of groups in the world with active MEIS programs increased with a factor of three from 1998 to 2006 and now numbers in the twenties. Many of these new groups are in industrial laboratories, and two commercial equipment manufacturers are marketing complete MEIS systems.

In depth profiling applications, the samples usually consist of an amorphous overlayer (film) of nm scale thickness on a single crystal substrate. The ion beam is here also aligned with a channelling direction in the substrate. This reduces the signal from the bulk of the sample, which may otherwise overlap and obscure the backscattered signal from light atoms in the film. The ion beam will now "see" all the atoms in the overlayer and also some of the atoms in the near interface region of the substrate. By analysing the shape of the signal from the different atoms, one can, using known parameters for straggling and energy loss, extract depth profiles, with in favourable cases, atomic layer resolution. The strength of the technique is not just the depth resolution in itself, but the fact that the beam can penetrate to a buried interface. Because interfaces play a very important role in many technologically important areas, there is a lot of potential for growth beyond the field of microelectronics, both in basic research and in applications. MEIS, initially developed as a basic research tool, has already proven its importance as a commercially significant analysis method.

2 Technical Aspects

The first high resolution MEIS measurements were made by the FOM group, using a large electrostatic ion energy analyzer [2]. The focus of the work with this instrument was surface structure determination. The main part of the instrument consisted of a cylindrical ion energy analyzer, the entrance and exit slits of which determined the ion energy resolution. The detector was rotated in the scattering plane so that different scattering geometries could be explored. Because only one detection angle and energy could be measured at one time, data accumulation times for a full investigation of a sample were quite long. The large size of the instrument also necessitated long bake-out times to reach ultra high vacuum. In spite of these limitations, it is fair to say that this instrument was absolutely essential in establishing the technique.

Data for multiple detection angles is crucial for surface structural work, which motivated Tromp et al. to construct a spectrometer (see chapter "Ion Spectrometers and Detectors") that detected many different angles at one time [3]. By analyzing the ion beam optics, they arrived at a design that allowed a $\sim 30°$ range of scattered ions (all in the scattering plane) to be collected at the same time. In order for proper focusing to take place, the deflecting metallic surfaces have to have a toroidal symmetry. Both an entrance and an exit slit were originally used. The signal from the ions passing through the exit slit was amplified by two multichannel plates (see chapter "Ion Spectrometers and Detectors"). The resulting electron pulse was detected by a position sensitive detector, so that the angular information could be recovered. This toroidal detector can be mounted on a table that can be rotated, so that different scattering geometries (all restricted to the scattering plane) can be investigated.

This instrument (Fig. 2), constructed in the early 1980s, forms the essential basis for the equipment used by a large part of the community today. The design was later modified by removing the exit slit and inserting a two-dimensional detector in its place [4]. Just as earlier, angular information was obtained in one coordinate, in the other, energy information. This obviously allowed a large reduction in the time to accumulate data. The toroidal design was commercialized by High Voltage Engineering in the Netherlands [5]. It features a resolution of $\Delta E/E$ of 4×10^{-3}. The commercial instrument has a mean bending radius for the ions of $10 \, \text{cm}$, with an entrance slit $5 \, \text{cm}$ away from the sample.

Other workers have explored different kinds of electrostatic detectors, which feature better energy resolution, but are often less flexible with regard to the scattering geometry (often constrained to a single geometry). They are generally much larger instruments. In particular, we would like to mention the electrostatic detectors built by Grötzschel et al. [6] and by Carstanjen et al. [7]. The latter instrument has a mean radius of $70 \, \text{cm}$ and operates at MeV energies.

Magnetic spectrometers are also very widespread. Kimura et al. have constructed a magnetic spectrometer with a bending radius of $15 \, \text{cm}$ [8]. It has been commercialized by Kobelco in Japan [9]. There is more limited flexibility in terms of scattering geometries, but one great feature of this instrument, aside from its very high resolution, is its small size (the commercial version has a footprint of only $2.1 \times 1.5 \, \text{m}$,

Fig. 2 Commercial toroidal
ion energy analyzer (High
Voltage Engineering,
Amersfoort, The
Netherlands). If the optional
2D position sensitive detector
is used, no exit slit is used
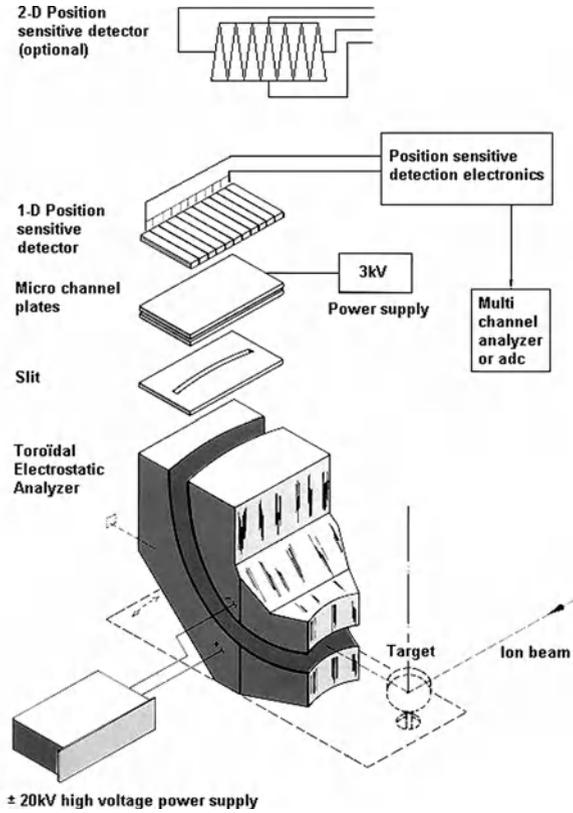
including a small vertical ion accelerator!). Due in part to its compact size it has
become very popular, particularly in industrial laboratories, where space is at a
premium.

An example of the impressive performance of this magnetic spectrometer is given
in Fig. 3 [8]. A sample of PbSe(111) was analyzed with 300 keV He ions. In this
orientation, this crystal consists of alternating layers of Pb and Se. Because Pb is
heavier than Se, the Pb signal is observed at higher energies. Note the clear separa-
tion between the first and second Pb layers, and the fact that the contributions from
layers deeper in the solid get broader because of straggling. From the width of the
leading Pb peak, one can estimate a depth resolution right at the surface of 0.01 nm.

Several groups have explored time of flight spectrometers. In the design by the
Vanderbilt University group the ion energy is determined by a start pulse from a foil
the backscattered ion passes through and by a stop pulse from a channel plate-type
detector [10]. Because of the unavoidable broadening of the ion energy due to the
foil, the energy resolution is only moderate. Another approach has been explored
by Kobayashi et al. [11]. In their analyzer, the incident ion beam is pulsed with
a pulse width of ~1.2 ns. This gives, with appropriate flight lengths, very com-
petitive energy resolution, at the cost of a low duty cycle (long data accumulation

Fig. 3 High resolution ion backscattering spectra of PbSe(111), obtained by Kimura et al. [8] using 300 keV He projectiles and a magnetic spectrometer. Notice the excellent depth resolution of the very first Pb peak
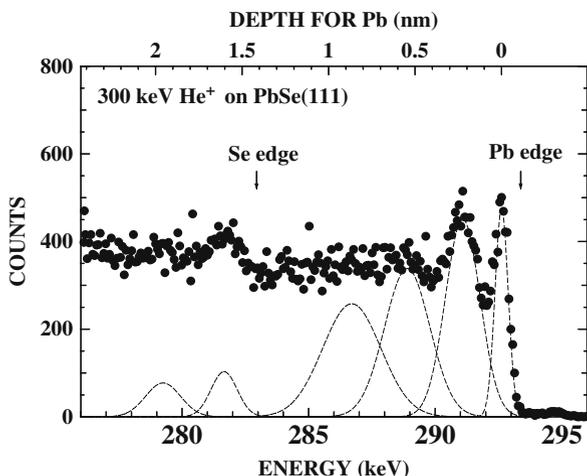
times). However, this spectrometer has other very interesting capabilities: It uses a set of large (120 mm diameter) two-dimensional channel plates as a detector and can therefore easily detect scattered ions away from the scattering plane. One then obtains a two-dimensional picture of the scattered intensity, a feature that none of the other detectors have. This is very important for studying surface structures that involve atomic displacements away from symmetry directions, and more generally for studying systems of low symmetry. Another advantage is that ions and neutrals are detected with the same efficiency; electrostatic and magnetic spectrometers only detect charged particles, and it is therefore very important to measure the charge fractions separately.

The accelerators for MEIS need to be quite stable and have a small energy spread, consistent with the needed energy resolution. In practice this means that single-ended machines are used, such as ion implanters with sufficiently low ΔE. A more detailed description of the relevant accelerator types is given elsewhere in this volume (see chapter "Ion Accelerators Used for Nanoscience").

3 Data Analysis

The theoretical background for MEIS is quite similar to that for RBS. The scattering potential is described by a screened Coulomb potential. As most ions are forward scattered from the ion cores, only those with impact parameters significantly smaller than the Thomas-Fermi screening length will contribute to the backscattering signal. The most common scattering potential used is the Moliere potential, which instead of the simple one-term Thomas-Fermi screening contains three terms with appropriate parameters determined from simple model calculations. A more detailed discussion of scattering potentials can be found in the chapter "Basis of Ion Scattering in Nanoscale Materials" of this volume.

Data interpretation for *surface structural work* is usually done using Monte Carlo simulations of the scattering yield. A possible model of the surface is used as input and its structural parameters are varied until a satisfactory fit (judged by a χ^2 type of analysis) to the (channelling and blocking) data is obtained. Data for different scattering geometries are sensitive to different structural parameters and should therefore be collected. As the probabilities of impingement on the subsurface atoms are strongly dependent on the vibrational parameters, these are also input parameters. Correlations between vibrations on adjacent atoms are often treated within a simple Debye model and anisotropies in the vibrational amplitudes have also been considered. As the backscattering probabilities are so small, the computer codes for these simulations have to be written in a way that optimizes computing times. The most common code is VEGAS, originating at the FOM.

As in the case of structural work, there is no data inversion procedure available for depth profiling. When interpreting thin film data, the sample is therefore considered to be divided into thin slices, each ∼0.3 nm thick. Beam propagation through this layer is modelled with an energy loss (taken from literature data, or, when those are not available, by simple averaging of similar compounds or elements) and assuming a Gaussian-shaped straggling function. The concentration of the different species is then varied by trial and error until an optimum fit is obtained.

For ultrathin layers, the nonstatistical nature of the ion-solid interaction has to be taken into account. The electronic structure of the atom becomes important, especially for heavy atoms, where the core electrons have large binding energies. Ionization of such electrons will give rise to an asymmetric backscattering peak and the assumption of Gaussian straggling therefore breaks down (see chapter "Basis of Ion Scattering in Nanoscale Materials"). This problem has recently been addressed by Pezzi et al., who have successfully implemented a model that accounts for the proper statistics of the small energy loss events and for an approximate electronic energy loss distribution during the backscattering event [12].

4 Research Examples

4.1 Surface Structure Determination

An example of surface structure determination is given in Fig. 4, from the work of Kido et al. [13]. They studied the surface termination of TiC(001). This surface contains in its ideal form equal numbers of Ti and C atoms. Although the surface layer has the same two-dimensional symmetry as the bulk, the C and the Ti atoms are not to be coplanar. The energy resolution used in this work was excellent, which allowed contributions from the first, second, and third layers to be deconvoluted. By analyzing the backscattered flux from the second layer Ti atoms and utilizing two different scattering geometries, they could get a very direct picture of the surface structure. In one geometry (Fig. 4a), they detected blocking by first layer Ti atoms. The surface blocking dip is clearly shifted towards larger angles, indicating an inward
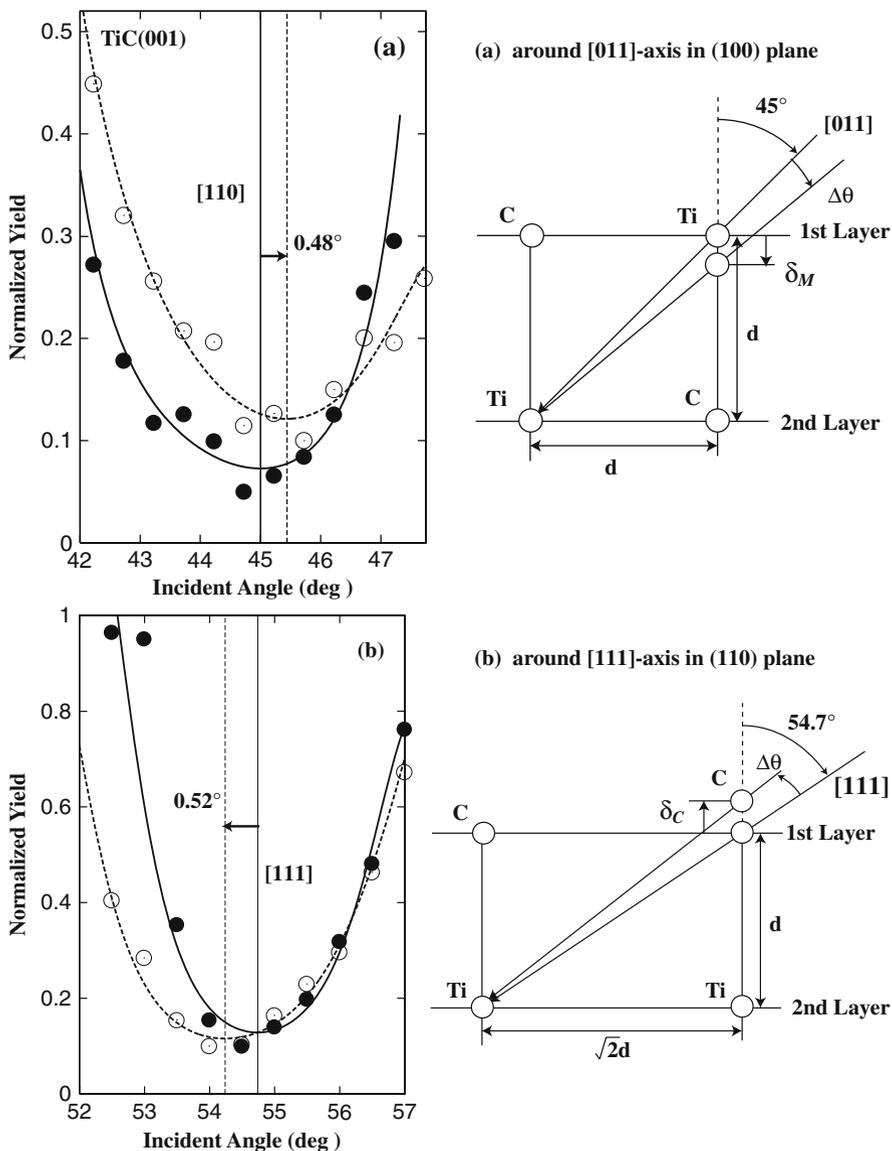
Fig. 4 Blocking curves obtained by Kido et al. [13] using $80\,keV$ He ions on TiC(001). In (**a**) results are shown for scattering around the [011] direction in a (100) plane, sensitive to the location of first layer Ti atoms. Scattering from the second layer is shown with *open circles* and from deeper layers with *full circles*. In (**b**) results are for scattering around a [111] axis in a (110) plane, sensitive to the location of the C atoms in the surface layer. An inspection of the relative signs of the position of the blocking dip in the surface scattering gives direct evidence for buckling of the surface layer

displacement of the Ti atoms in the first layer. On the other hand, the geometry in Fig. 4b, sensitive to blocking by the carbon atoms in the surface, is displaced in the opposite direction, implying that the C atoms have moved outwards. A detailed evaluation of the data revealed that this rumpling has an amplitude of 3.5% of the bulk interlayer spacing. No evidence was found for a significant overall surface layer relaxation.

Surface structural work continues to be an important area in MEIS. There has been much recent progress on the structure of metallic systems, including both alloys [14] and adsorbates [15, 16], as well as on silicates [17, 18], yielding structural information not easily obtainable by other means. Another area that cannot be discussed here because of space limitations involves the path-breaking experiments on nanoscale surface-induced melting observations using MEIS [19].

4.2 Depth Profiling

RBS is a well-established method for depth profiling of thin films with a resolution often in the ∼10 nm range. MEIS carries this one step further into the sub-nm region. Aside from the high energy resolution, the fact that the electronic stopping power for light projectiles reaches a maximum at typical MEIS energies also contributes to this improvement. The main scientific problem studied has been ultrathin films on semiconductors, primarily Si. The driving force for this is the decreasing size of microelectronic components (see chapter "High Speed Electronics"). In particular, the thickness of the gate oxide in CMOS is now on the order of one nm. An atomic layer understanding of the formation mechanisms and the thermal stability of this oxide is therefore important. Historically, this oxide was silicon oxide, a very stable, defect-free material. However, when the thickness of this layer reached the nm range, leakage currents became far too large and new materials, which can form thicker barriers, need to be explored. In order to have the same electrical control, these layers need to have a higher dielectric constant (κ or k). Suitable high-k materials are ZrO_2 and HfO_2 and their silicates, which have good thermal stability on silicon under common processing conditions (see chapter "High Speed Electronics").

4.2.1 Oxidation of Silicon—The Use of Isotopes

Oxidation of silicon is one of the most studied topics in materials science and a vast literature exists. Here, we will discuss some of the contributions MEIS has made to that subject. In order to elucidate the formation mechanisms of oxides, isotope experiments are very useful. Isotope marking allows us to study not just the structure of thin films, but also dynamic processes. The two different oxygen isotopes are easily resolved. In Fig. 5 we show MEIS spectra (only the oxygen region is shown) for a 4.5 nm silicon oxide film grown on Si(100) [20, 21]. The film was grown in ^{16}O and then reoxidized in ^{18}O. The original film (filled symbols) shows a sharp leading edge and a broadened lower edge. This broadening is in part
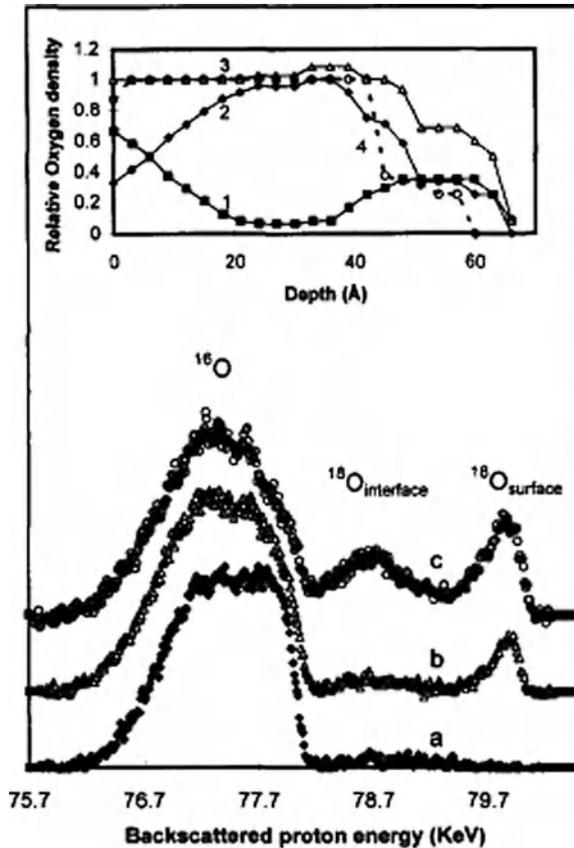
Fig. 5 Oxygen spectra for (**a**) a 6 nm oxide grown on Si(100); (**b**) the same sample reoxidized in SO, at 1 Torr. 800 °C for 3 h; and (**c**) following further reoxidation for 21 h. The insert shows oxygen depth profiles in the oxide determined by model simulation of the spectrum (**c**)

due to straggling, but also to a concentration gradient at the oxide-silicon interface. When the sample is reoxidized in ^{18}O, two peaks appear at higher (due to the higher atomic mass) backscattering energies. The energy of the leading peak corresponds to ions backscattered from ^{18}O atoms in the surface. There is a reduction of the backscattered intensity from the lighter isotope at the surface. One can therefore conclude that the incoming ^{18}O exchanges with the ^{16}O already incorporated in the film. The origin of this exchange reaction is not fully understood, but it is influenced by impurities in the sample surface. At lower energies in the ^{18}O atom region a second peak is observed, corresponding to backscattering from atoms close to the interface. This is a growth reaction. The fact that these two peaks are well resolved with little intensity between them implies that oxygen travels through the film as a molecule and dissociates and reacts near the interface. At the same time, the low energy edge of the ^{16}O distribution moves further down, which implies a complicated mixing behavior near the substrate.

4.2.2 Measurement of Strain at the Si/SiO$_2$ Interface

The structure of the SiO$_2$/Si(100) interface is obviously of central importance in microelectronics. In order to get detailed structural information about the interface, Kim et al. have measured the angular distribution of backscattered ions for atoms located at the transition region [22]. Some of their results are shown in Fig. 6. A very well developed and clear blocking dip is observed for backscattering energies corresponding to the bulk (substrate) Si crystal. The deep blocking dips in the bulk Si change to shallow dips in the transition layer, which means that the crystallinity in the transition layer is lower than in the substrate. These blocking dips were taken at the SiO$_2$ side of the interfaces. A very interesting observation is the shift of the <111> blocking dips in the interface layer to higher scattering angles with respect to the dips in the bulk Si. The blocking dip shift for thermal oxides is smaller than those for the ion beam-grown oxides. Based on a simple trigonometric estimation, the angular shifts of <111> dips in the thermal oxide and the ion beam oxide, respectively, can be converted to 0.0011 and 0.0038 nm vertical displacements of the Si atoms, which correspond to 0.96 and 2.8% vertical strains (expansion) of the Si lattices in the interface. This can be understood in terms of stress because of the volume expansion from Si oxidation. It was later shown that oxides grown in ozone do not show a detectable angular shift, implying that the ozone oxide is

Fig. 6 The blocking dips of scattering from Si in the transition layers of a thermal oxide and ion beam oxides before and after annealing [22]. The shift in the blocking dip position indicates strain in the substrate
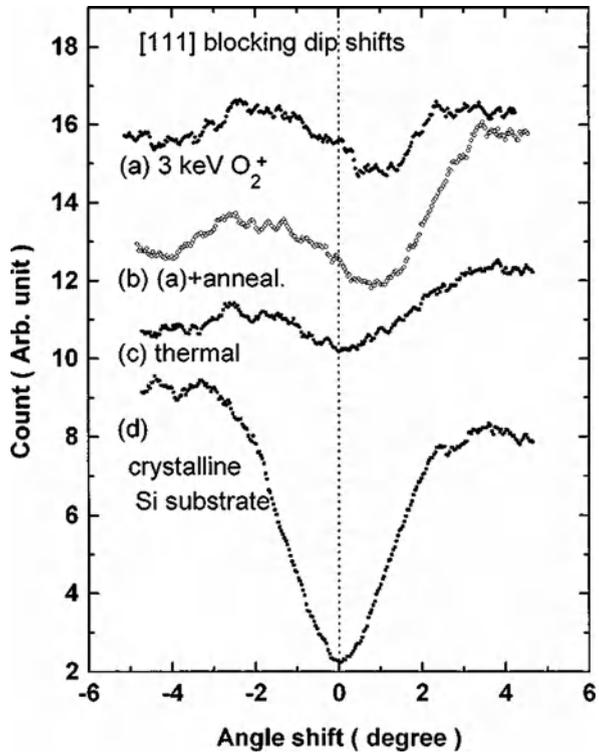
homogenous [23]. These results have been correlated with electrical measurements of failure rates as a function of breakdown voltage. Nakajima et al. have used the same methodology to do strain profiling at the $HfO_2/Si(100)$ interface [24]. The blocking dips were here shifted in the opposite direction, indicating the existence of *compressive* vertical strain. The maximum strain was about 1%, decreasing rapidly with increasing depth. The observed strain may have a large effect on the carrier mobility in HfO_2-based MOSFETs.

4.2.3 High-k Materials on Silicon

When high-k dielectrics are deposited on Si, normally a thin beneficial SiO_2 layer exists between the metal oxide and the Si substrate. One needs to understand if the metal ions remain segregated into a separate layer, and how this behavior changes for a composite film. In the case considered here, the composite layer consists of an HfO_2 film, with a thin capping layer of BaO. A great deal of development work has been invested in HfO_2 as a gate dielectric, and most of the electrical properties are at the stage where it is a viable replacement for SiO_2. However, the threshold for device turn-on is not practical for CMOS applications. Thus, there is interest in modifying the threshold voltage without sacrificing the mobility and reliability that have been attained for HfO_2. If a capping layer can be added to modify the dielectric stack with minimal impact on device performance, that would enable its use in CMOS. Group IIa and IIIa elements (La, Mg, Ba, etc.) have been identified as candidates for this role, since they lower the threshold voltage for nFETs.

Annealing causes intermixing of BaO layers deposited directly on SiO_2. This can be seen in Fig. 7a, which shows MEIS spectra [25] from 0.26 nm of BaO on 2.7 nm of SiO_2. Upon annealing, the Ba peak is seen to broaden, because of dilution with SiO_2. After the sample has been heated to $900\,°C$, the Ba peak has expanded until it has nearly the same width as the oxygen peak. The inset in Fig. 7a shows the Ba and oxygen depth distributions plotted on the same scale: The Ba ions have spread throughout the dielectric. The behavior for La is quite different; after annealing, the La ions are confined to the top region of the sample (Fig. 7b). The La peak does broaden, but it still remains much narrower than the oxygen peak. Evidently, a silicate phase nucleates for La, and this occupies the upper part of the sample. For Ba, a silicate phase does not nucleate, and instead the Ba dissolves into the SiO_2.

Unfortunately, both Ba and Hf are heavy elements, and it is difficult to achieve sufficient mass resolution with conventional proton beams to clearly discriminate between the two. This has necessitated the use of a $200\,keV$ He^+ beam at a large scattering angle. Although multiple charge states are a complication with He^+ projectiles, we can get a qualitative measure of the depth distribution without undue complications from the final ionization state. Figure 8 shows the Ba portion of the spectra for a sample before and after $1000\,°C$ rapid thermal anneal. Initially, the Ba is confined to a narrow surface region. After annealing, the Ba diffused into the dielectric, with a pileup at the HfO_2/SiO_2 interface. Two simulations are shown, one for Ba distributed throughout the SiO_2, the other for Ba confined to a narrow region of $BaSi_2O_5$ at the HfO_2/SiO_2 interface. The conclusion is that the

Fig. 7 MEIS spectra from 0.26 nm of BaO on 2.7 nm of SiO_2, as deposited and after annealing. Upon annealing, the Ba peak is seen to broaden. After the sample has been heated to 900 °C, the Ba peak has expanded until it has nearly the same width as the oxygen peak. The inset in (**a**) shows the Ba and oxygen depth distributions plotted on the same horizontal scale. The spectra have been scaled vertically taking cross-sections into account. The results in (**b**) for La-containing oxides are very different. The vertical scale is in arbitrary units
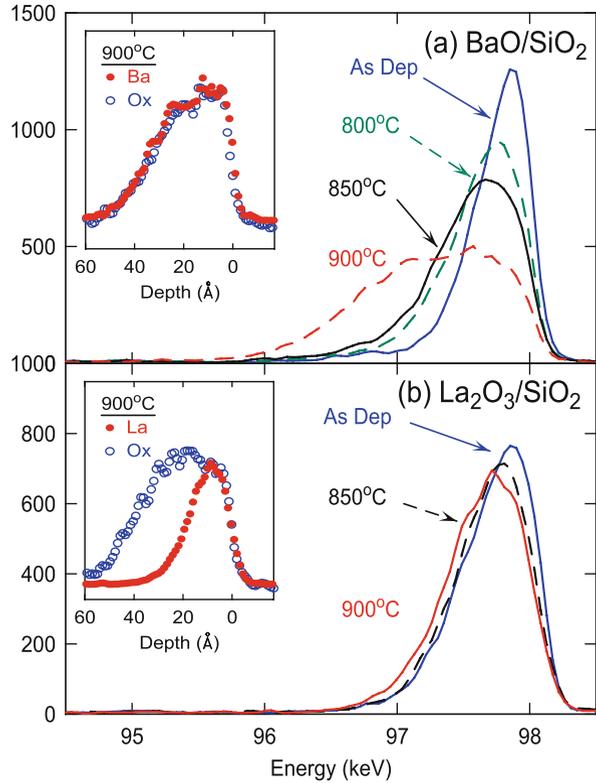
Fig. 8 MEIS spectra of a composite $BaO/HfO_2/SiO_2$ film. A dramatic shift of the Ba peak is observed as the sample is annealed. The vertical scale is in arbitrary units
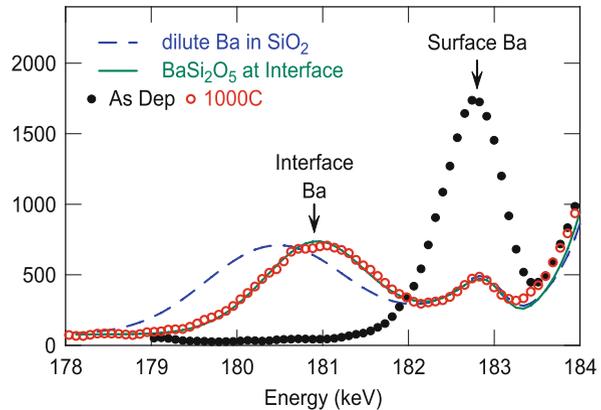
HfO_2/SiO_2 interface acts as a nucleation center for Ba silicate. Note that the Ba distribution is not a simple gradient, but that the concentration has a minimum in the HfO_2. Also, the majority of Ba atoms diffuse to the interface, indicating that silicate formation is highly efficient in this system. A final point to consider is that migration away from the metal/dielectric interface increases the dipole moment of the

Ba ions, thereby strengthening their effectiveness in altering the threshold voltage offset.

4.2.4 Epitaxial Oxides on Silicon

Epitaxial oxides on silicon have attracted a great deal of attention, originally for a possible use in high-k layers in devices, but more generally as a way for epitaxial integration of new functionality on a silicon platform. One such system is $SrTiO_3$ on Si(100), which in spite of a 2% lattice mismatch can be grown epitaxially. The growth of such a structure is complicated. It involves first depositing Sr on the clean Si substrate at elevated temperatures to form a Sr silicate. Further Sr is deposited and oxidized to form SrO. This is followed by TiO_2 deposition, annealing, recrystallization, etc. In Fig. 9, we show results of MEIS studies of two such films [26]. As explained above, because of channeling in the epitaxial structure, ion backscattering will only be observed from the surface layer(s) *and* from imperfections in the crystalline order, especially at the interface. The data in (b) clearly show two distinct oxygen-related peaks, one from the surface and one from the interface region. This is evidence for a distorted region at the interface. Likewise, there is backscattering from Ti in the interface region. A detailed analysis shows that instead of having Sr atoms closest to the substrate, there is here a Ti silicide—implying that the position of the metal atoms has been reversed. For a lower annealing temperature (Fig. 9a), one instead obtains a thin layer of Sr silicide. The ability to obtain such
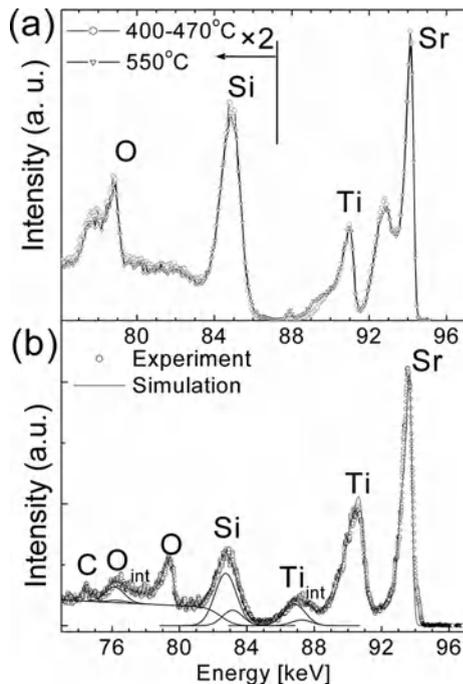
Fig. 9 (**a**) MEIS energy distributions for a 3.5 nm thick $SrTiO_3$ films on Si(001) recrystallized at $450\,°C$ and $550\,°C$, and (**b**) a 7.8 nm thick film recrystallized at $550\,°C$. Decomposition of the interface peak is demonstrated for the thicker film

a very detailed picture of a buried interface (the films in Fig. 9 are 3.5 and 7.8 nm thick) is an important strength of MEIS. This information can now be fed back to crystal growers and result in refinements of the growing methodology.

References

1. J.F. van der Veen. Surf. Sci. Rep. **5**, 199–288 (1985)
2. W.C. Turkenburg, W. Soszka, F.W. Saris, H.H. Kersten, and B.G. Colenbrander. Nucl. Instrum. Methods **132**, 587–602 (1976)
3. R.M. Tromp, H.H. Kersten, E. Granneman, F.W. Saris, R. Koudijs, and W.J. Kilsdonk. Nucl. Instrum. Methods B **232**, 155–66 (1984)
4. R.M. Tromp, M. Copel, M.C. Reuter, M. Horn von Hoegen, J. Speidell, and R. Koudijs. Rev. Sci. Instr. **62**, 2679–83 (1991)
5. http://www.highvolteng.com/
6. R. Grötzschel. (private communication)
7. H.D. Carstanjen. Nucl. Instrum. Methods. B. **136–138**, 1183–90 (1998)
8. K. Kimura, K. Ohshima, K. Nakajima, Y. Fujii, M. Mannami, and H.J. Gossmann. Nucl. Instrum. Methods **B99**, 472–5 (1995)
9. http://www.kobelco.co.jp/technobook/p238_e.htm
10. M.H. Mendenhall and R.A. Weller. Nucl. Instrum. Methods. **B59–B60**, 120–3 (1991)
11. T. Kobayashi. Nucl. Instrum. Methods. **249**, 266–9 (2006)
12. R.P. Pezzi, P.L. Grande, M. Copel, G. Schiwietz, C. Krug, and I.J.R. Baumvol. Surf. Sci. **601**, 5559–70 (2007)
13. Y. Kido, T. Nishimura, Y. Hoshino, S. Otani, and R. Souda. Phys. Rev. B **61**, 1748–51 (2000)
14. D. Brown, P.D. Quinn, D.P. Woodruff, T.C.Q. Noakes, and P. Bailey. Surf. Sci. **497**, 1–12 (2002)
15. T.G. Owens, T.E. Jones, T.C.Q. Noakes, P. Bailey, and C.J. Baddeley. J. Phys. Chem. B **110**, 21152–60 (2006)
16. M.A. Munoz-Marquez, G.S. Parkinson, P.D. Quinn, M.J. Gladys, R.E. Tanner, D.P. Woodruff, T.C.Q. Noakes, and P. Bailey. Surf. Sci. **582**, 97–109 (2005)
17. T.J. Wood, C. Bonet, T.C.Q. Noakes, P. Bailey, and S.P. Tear. Phys. Rev. B **73**, 235405–1 (2006)
18. D.J. Spence, T.C.Q. Noakes, P. Bailey, and S.P. Tear. Surf. Sci. **512**, 61–6 (2002)
19. J.W.M. Frenken and J.F. van der Veen. Phys. Rev. Lett. **54**, 134–41 (1985)
20. E.P. Gusev, H.C. Lu, T. Gustafsson, and E. Garfunkel. Phys. Rev. B **52**, 1759–75 (1995)
21. E.P. Gusev, H.C. Lu, T. Gustafsson, and E. Garfunkel. Appl. Surf. Sci. **104–105**, 329–34 (1996)
22. Y.P. Kim, S.K. Choi, H.K. Kim, and D.W. Moon. Appl. Phys Lett. **71**, 3504–6 (1997)
23. A. Kurokawa, K. Nakamura, S. Ichimura, and D.W. Moon. Appl. Phys. Lett. **76**, 493–5 (2000)
24. K. Nakajima, S. Joumori, M. Suzuki, K. Kimura, T. Osipowicz, K.L. Tok, J.Z. Zheng, A. See, and B.C. Zhang. Appl. Phys. Lett. **83**, 296–8 (2003)
25. M. Copel. Appl. Phys. Lett. **92**, 152909 (2008)
26. L.V. Goncharova, D.G. Starodub, E. Garfunkel, T. Gustafsson, V. Vaithyanathan, J. Lettieri, and D.G. Schlom. J. Appl. Phys. **100**, 14912–1 (2006)

Box 5: Surface Crystallography Terminology

Harry J. Whitlow and Sachiko T. Nakagawa

1 Introduction

The crystalline nature of the surface differs from the bulk because atoms on the surface experience a different force field due to unterminated bonds, oxidation by adatoms etc. [1]. Free energy minimisation leads to *reconstruction* of the surface layer from the bulk by formation of dimers and displacement of atoms from their normal sites. This is illustrated in Fig. 1 where the surface atoms are displaced by d_1 and d_2 in the vertical and lateral directions, respectively.
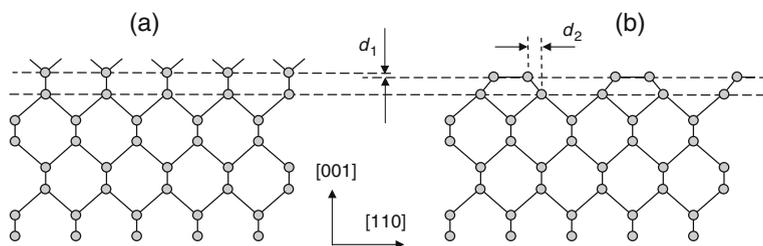
Fig. 1 Schematic illustration of surface reconstruction by dimer formation of the (001) surface of a diamond lattice. (**a**) Bulk lattice without dimer formation, (**b**) surface reconstruction by formation of dimers. d_1 and d_2 denote the vertical and lateral displacement of the surface superlattice atoms

It is usual to describe an ordered superlattice of surface atoms relative to the 2-dimensional crystal lattice of the substrate surface. The symmetry of 2-dimensional crystals is based on a combination of rotational, inversion and mirror symmetry operations perpendicular to the surface plane. This gives 5 Bravais lattices and 10 point groups which combine with screw and translation symmetry operations yielding 17 possible plane groups [2]. Defining superlattice lattice vectors $\mathbf{b_1}$, $\mathbf{b_2}$ in terms of the substrate lattice vectors $\mathbf{a_1}$, $\mathbf{a_2}$ as, $\mathbf{b_1} = m\mathbf{a_1}$, $\mathbf{b_2} = n\mathbf{a_2}$. This may be combined with a rotation R of the superlattice vectors $\mathbf{b_{1,2}}$ with respect to the substrate

H.J. Whitlow (✉)
Department of Physics, University of Jyväskylä, Jyväskylä FIN-40014, Finland
e-mail: harry.j.whitlow@jyu.fi

$\mathbf{a_{1,2}}$. Figure 2 illustrates the notation which, for a superlattice on a M$\{hkl\}$ substrate surface is written in shorthand as: M$\{hkl\}$ $(m \times n)$ R and M$\{hkl\}$ c $(m \times n)$ R, for primitive (Fig. 2(a), (b), (d)) and centred (Fig. 2(c)) superlattices, respectively. R denotes a rotation of the superlattice relative to the substrate.
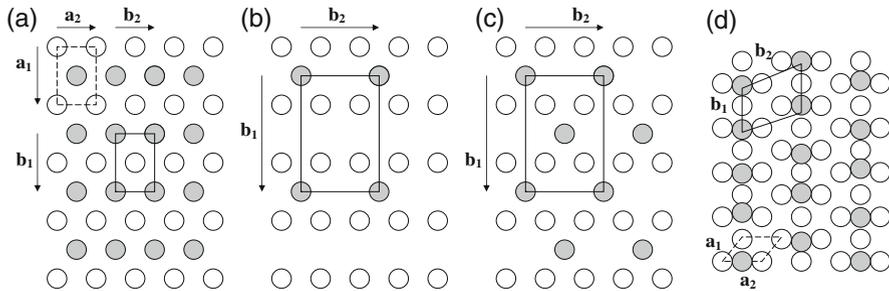
Fig. 2 Examples of notation for a superlattice (*grey atoms*) on a generic elemental crystal substrate M (*white atoms*). (**a**) M$\{100\}$ (1×1), (**b**) M$\{100\}$ (2×2), (**c**) M$\{100\}$ c (2×2), (**d**) M$\{111\}$ (2×2) $30°$

Surface defect nomenclature is illustrated in Fig. 3. Vicinal planes where $\{hkl\}$ are large are made up of low-index plane terraces with monolayer ledges (1-dimensional defects). Adatoms and atomic vacancies on the ledge and terrace plane constitute zero-dimensional defects.
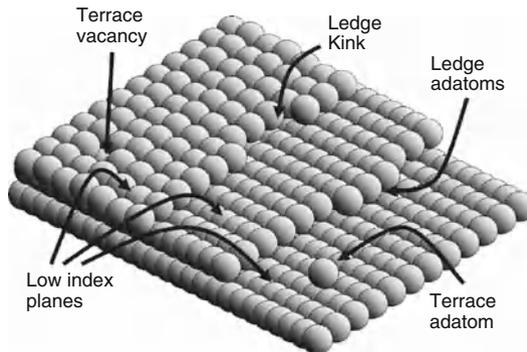
Fig. 3 Illustration of a vicinal surface showing various surface defects

References

1. H. Lüth: *Solid Surfaces, Interfaces and Thin Films*, 4th edn (Springer, Berlin Heidelberg New York, 2001) pp. 77–94
2. H. Burzlaff, B. Gruber, P.M. de Wolf, and H. Zimmermann, In: *International Tables for Crystallography*, vol A, ed Th. Hahn, 5th edn (International Union for Crystallography by Kluwer, Dordrecht, 2002)

Thin Film Characterisation Using MeV Ion Beams

Timo Sajavaara and Kai Arstila

This chapter focuses on the characterisation of very thin films having thicknesses from a few nanometres to tens of nanometres. The driving force for the ion beam analysis community has mostly been the rapid development of microelectronics — all the elements in new thin SiO_2 replacing dielectrics, diffusion barriers, and silicide contacts need to be analysed with a depth resolution even better than a nanometre. This together with new film deposition techniques like atomic layer deposition (ALD) [1] have given a push to the ion beam analysis community to develop new and better techniques using energetic (>0.5 MeV) ion beams.

The main techniques in this energy region are the Rutherford backscattering spectrometry (RBS), the elastic recoil detection analysis (ERDA), and the nuclear reaction analysis (NRA). The basic concept in all these techniques is similar: the sample under characterisation is irradiated with energetic ions, and emitted particles or electromagnetic radiation are detected. The depth information is based on energy losses of ingoing and outcoming particles. In addition to depth profiling in nanoscale, one can also scan a focused beam on a sample surface and detect emitted characteristic X-rays with a spatial resolution on the order of hundreds of nanometres using particle induced X-ray emission (PIXE) technique. In Table 1 these different techniques are highlighted, with their common performance figures. Among these techniques, high-resolution RBS (HR-RBS) and heavy-ion ERDA (HI-ERDA) are the most versatile and the most used ones for nanometre depth profiling, so the emphasis of this chapter is mainly on these two techniques.

1 Conventional Rutherford Backscattering Spectrometry

Although the first backscattering experiments with energetic alpha particles were done almost one hundred years ago, in the years 1909–1911 by Ernest Rutherford and his students, RBS was developed to its still most commonly used form during

---

T. Sajavaara (✉)

Department of Physics, P.O. Box 35, University of Jyväskylä, Jyväskylä FIN-40014, Finland
e-mail: timo.sajavaara@jyu.fi

Table 1 Different IBA techniques which can be used in the characterisation of thin films

Technique	Incident beam	Beam energy (MeV)	Elemental sensitivity	Surface depth resolution (nm)	Limitations, comments
RBS	^{4}He	1–2	> He, > substrate	10–20	Simple setup, standard software
ERDA	^{4}He	1–10	^{1}H, ^{2}H	≈100	Simple setup, absorber foil
HR-RBS	^{4}He, Li, C	1–2	> He, > substrate	<1	Magnetic or TOF spectrometer
NRA	^{2}H, ^{15}N, ^{19}F	0.3–10	H, C, N, O	1–10	Isotope selective, good depth resolution at large depth
HI-ERDA	Cl, Cu, Br, I	5–100	All elements	2–10	TOF or gas detector, multi-parameter measurement

the 1960s. The main reason for this step was the development of semiconductor-based detector technology, which brought this powerful analysis technique to all accelerator laboratories.

In RBS, only the backscattered probing ion is detected, and the depth in which the scattering occurred can be calculated by means of scattering kinematics and stopping forces (see chapter "Basis of Ion Scattering in Nanoscale Materials"). For a standard measurement, 1.0–2.5 MeV He ions are used; and in a typical measurement geometry the energy detector is at an angle of 150–170°, and the sample surface is perpendicular to the direction of the incoming beam. This measurement geometry has its advantages: (i) no need for particle identification because only the backscattered incident ions are detected, (ii) the resolution between different masses is at its maximum at a 180° scattering angle, (iii) a small discrepancy in the detector angle has a very small effect on the analysis, and (iv) sample surface roughness effects are at their minimum. A typical RBS measurement from a 390 nm thick WO_3 film deposited on glassy carbon is shown in Fig. 1.

When a homogeneous thin film, with all the elements detectable is measured, atomic concentrations can be calculated from the backscattering yields in the energy spectrum according to the Rutherford scattering cross-sections (see chapter "Basis of Ion scattering in Nanoscale Materials"). Moreover, the stopping cross-sections of many materials for most commonly used hydrogen and helium ions are experimentally well known. The major drawback of RBS is its limited sensitivity to detect and quantify low masses (<20 amu) and, in general, elements lighter than those in substrate. This is a severe limitation of RBS for characterisation of thin films grown on, for instance, germanium. In the thin film process development, the lightest impurity, hydrogen, is impossible to measure by RBS, and the sensitivity for carbon and oxygen is very limited. The depth resolution of about 20 nm at the surface is also a drawback of the technique. On the other hand, by increasing energy, composition information can be obtained from a depth of several micrometres, which is not possible for all the depth profiling techniques.
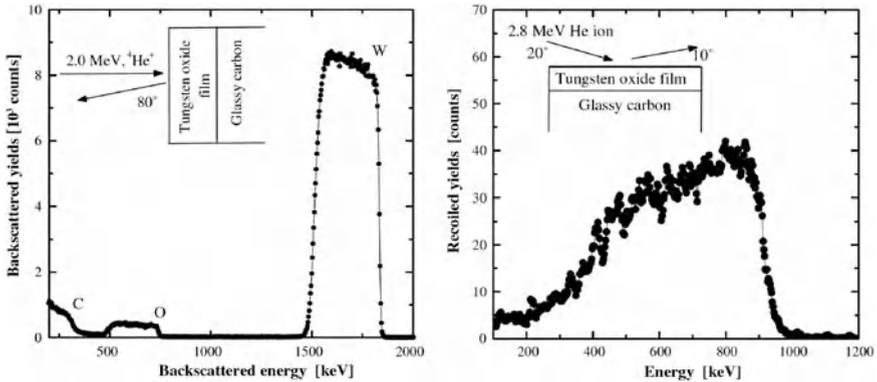
Fig. 1 Typical RBS spectrum (*left*) obtained by using 2.0 MeV He ions and typical ERDA spectrum (*right*) obtained by using 2.8 MeV He ions for an oriented tungsten oxide film deposited on glassy carbon at a substrate temperature of 600 °C. The O/W ratio was determined to be 3.00±0.05 and hydrogen concentration 5.8 at.% from RBS and ERDA measurements, respectively [2]

2 Conventional Elastic Recoil Detection Analysis with Range Foils

In contrast to RBS, forward-scattered sample atoms are detected in ERDA instead of incident ions. In the first measurement configurations, a simple charged particle detector, similar to those used in RBS, was set to a forward angle (typically 30–45° with respect to the in-going ion beam) and protected by micrometre-thick Al or poly(ethylene terephthalate) (PET) film [3, 4]. This was to prevent incident ions reaching the detector and to allow only the lightest recoils with smallest energy loss in the absorber (normally hydrogen isotopes) to do so. Due to small stopping forces for both light incident ion and light recoil in the sample and due to energy straggling in the absorber foil, only modest depth resolution of several tens of nanometres can be achieved. Furthermore, reference samples with known composition need to be used for quantified analysis. An example of a conventional ERDA spectrum is shown in Fig. 1. In this measurement, a 12 μm thick Al-foil was placed in front of the detector to prevent incident He ions scattering to the detector.

3 Nuclear Reaction Analysis

In NRA, the Coulomb barrier between two colliding nuclei is penetrated, the nuclei interact with each other, and reaction products, particles or γ-rays, are detected. The most often used techniques are ^{2}H beam for ^{12}C [5], ^{14}N [6] and ^{16}O [7] profiling, and ^{15}N [8] and ^{19}F for H profiling.

NRA is considered to be the most sensitive one among depth profiling ion beam analysis techniques. Concentrations down to ppm level can be measured for some elements and, especially for particle-gamma reactions with a sharp resonance, even

nanometre depth resolution can be achieved [9]. Yet the determination of the absolute concentrations requires the use of standards, and detailed measurements using narrow resonances are normally time-consuming and sensitive samples also suffer from high fluences. Furthermore, additional RBS or HI-ERDA measurements are normally required if all elements in the thin film need to be analysed.

An additional high energy ion beam technique is particle- or proton-induced X-ray emission (PIXE) [10], in which ions in the material slow down and excite atoms, which then can emit characteristic X-rays that can be detected. Because of its physical working principle, this technique has basically no depth resolution, but when scanned over a sample with a highly focused nanobeam, it can be considered to be a nanoscience technique. The minimum beam current requirement of 100 pA for doing PIXE with a scanning beam limits the achievable spatial resolution to around 400 nm [11].

4 High-Resolution RBS

The depth resolution of the RBS technique can be greatly improved by means of detectors with far better energy resolution than silicon solid state detectors. The use of these detectors will bring the resolution even to the subnanometre range, but with the cost of more complicated and expensive measurement instruments and data-handling procedures. The two most common ways to do high-resolution RBS (HR-RBS) are to use a magnetic spectrometer [12–14] or a time-of-flight (TOF) spectrometer [15, 16].

4.1 RBS with Magnetic and Time-of-Flight Spectrometers

The basic magnetic spectrometer used in RBS analysis consists of a 90° magnet and a position sensitive detector. In most cases the energy of the scattered particle is not directly measured, but it can be deduced from the particle position after using the energy analysing magnet. The position is usually measured by a position sensitive microchannel plate (MCP) assembly. The setup in Kyoto, Japan, by Prof. Kimura is shown in Fig. 2. Even with a small 90° magnet with a bending radius of 200 mm, an energy resolution on the order of 1 keV can be obtained for scattered 400 keV He ions. This instrument, including a 500 kV accelerator and full analysis software, has been commercialised by the Japanese company Kobe Steel.

Although the spectrometers are normally made movable, it is common to position the magnetic spectrometer to a forward angle of 60–90° with respect to the incident beam (Fig. 3). This is motivated by several reasons. Firstly, the solid angle of the detector is generally quite small, on the order of 0.5 msr, and the forward scattering angle with higher scattering cross-section decreases the average measurement time. Secondly, the energy window of the spectrometer, without changing the magnet current, is between 10% and 30% of the central energy. In the backscattering angles normally used in the conventional RBS measurements, events originating
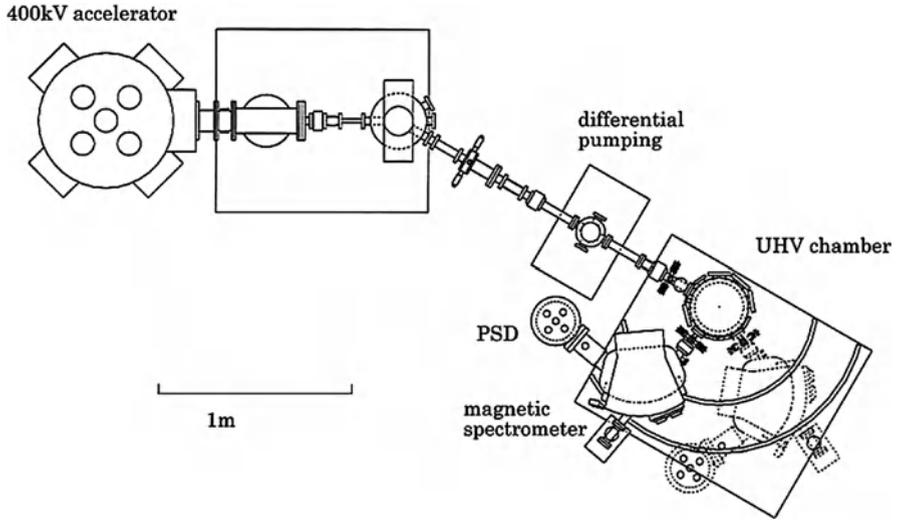
Fig. 2 A compact high-resolution RBS system consisting of a simple magnetic spectrometer and a small accelerator in Kyoto, Japan [17]

from heavier thin film elements and, for instance, silicon substrate could not be collected at the same time because of the limited energy window, making the analysis less straightforward. The worse mass-separation is compensated for by much better energy resolution. Thirdly, the depth resolution is improved because the forward scattering angle increases the path length of both ingoing and outcoming ions in the sample.

A major limitation of a magnetic spectrometer is the fact that ions can have several different charge states when they exit the sample because of the depth of scattering or the surface material [18]. Only ions with a certain charge state should
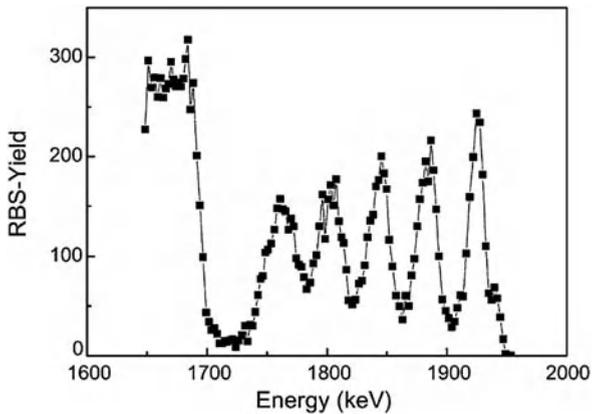
Fig. 3 Spectrum of a multilayer $5 \times (0.57\,nm\ Ta_{0.1}Si_{0.9}/5.2\,nm\ Si)$ on Si: RBS spectrum measured with 2 MeV C-ions at a scattering angle of 35.5° using a magnetic spectrometer [14]

be collected at the same time, and in the analysis it is mostly assumed that the charge state distributions do not vary between scatterings taking place in different depths. In the case of He, only charge states 0–2^+ are possible, but for heavy incident ions the distribution is greater and therefore they are not generally used for analysis.

An interesting extension to the technique is to use the same forward angle magnetic spectrometer for hydrogen ERDA. The resulting depth resolution is far better than the one achieved by range foils [19].

The other way to outperform silicon detectors is to measure the energy or, to be more precise, the velocity spectrum for scattered ions by means of a time-of-flight telescope [15, 16, 20]. This telescope consists of two timing detectors (explained in more detail later, in the heavy-ion-ERDA section), and with known path length, the velocity for detected ions can measured, which then can be converted to the energy spectrum. The performance of a TOF detector is not better than a silicon solid state detector for He ions in the energy range above 1 MeV. For lower energy He ions, however, and especially for heavier incident ions a TOF detector clearly outperforms silicon detectors. The performance of a state-of-the-art TOF telescope for low energy He ions is shown in Fig. 4 [21]. In contrast to magnetic spectrometers, this detector type is not sensitive to charge state distributions of ejected particles, and the whole energy range can be collected during a single measurement.

One interesting variation of a TOF-RBS-spectrometer is to apply a TOF-ERDA telescope in forward angles using light incident ions like C or O for high sensitivity and good depth resolution analysis [22].
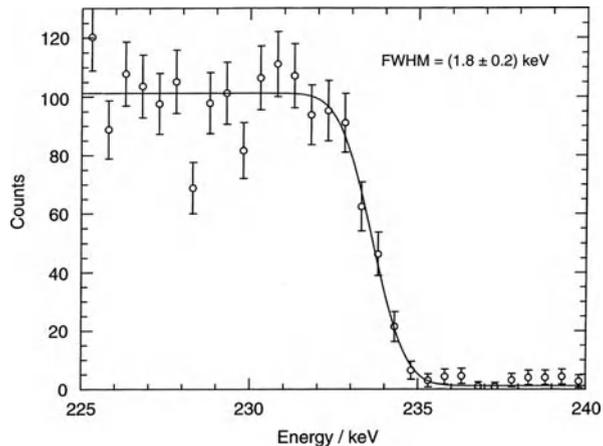
Fig. 4 The TOF-RBS example shows a spectrum obtained with a 1 MeV He beam backscattered from a carbon bulk sample [21]. Start foil was $0.6 \, \mu g/cm^2$ diamond-like carbon foil. TOF telescope was located at $175°$, flight path was 1.4 m, and timing resolution 200 ps

5 Heavy-Ion Elastic Recoil Detection Analysis

When a particle detector is positioned to a forward angle in the ion beam measurement of a thin film, it becomes possible to detect sample atoms recoiled by

the primary ions. By quantifying the energy and the yield of the recoils, one can determine the composition of a thin film sample for all atomic species. Since the kinematics of a forward angle measurement are not limited by the requirement for backscattering, the mass of the primary ion can be chosen freely.

Since the cross-section of the recoil yield increases to the power of four (for recoils much lighter than the beam) for the atomic number of the primary ion, the choice is usually a very heavy primary ion. Also, the stopping force for heavy ions in target materials is much higher than for light ions, which results in higher depth resolution. The use of heavy ions requires an ion source capable of creating beams from heavy ions, which are most often based on the sputtering of solid source materials. This ends up in a common choice of ion beams from halogens Cl, Br, or I. Also transition metals like Cu, Ag, and Au are regularly used.

A primary ion which is heavier than the matrix material of the substrate (in the case of thin films this is very often silicon) allows the choice of a detector angle larger than the critical angle, where primary ions can scatter from substrate material. The detector angle is thus usually chosen to be larger than typical critical angles, but not too high to reduce recoil energies too much. In practise, in most of the HI-ERDA detector systems the detector angle is fixed to a single value which is between $35°$ and $45°$.

Since all different sample atom species, and also the primary ions in some cases, are scattered to the forward detector, it is necessary to have a mass or element selective detector to allow reasonable data analysis. This requirement, together with the need for relevant electronics and data acquisition systems and a heavy ion source, forms the largest obstacle to a more common use of HI-ERDA in the analysis of thin films and nanomaterials compared to more common IBA methods such as RBS.

Most common selection mechanisms with HI-ERDA detectors are based on the measurement of time-of-flight (velocity) in coincidence with energy for each particle, or dE–E type of detectors based usually on the slowing down and ionisation in gas. Magnetic spectrometers are in some cases also used for HI-ERDA, but they are not only complicated and expensive, but they also have difficulties in determining total yields because of the different charge state distributions of each recoiled target atom species.

5.1 Detectors for HI-ERDA Measurements

In TOF detectors, the timing signal is usually obtained from multiplication of secondary electrons by MCPs emitted from a thin film due to the transmission of an ion. A simultaneous energy signal is usually obtained with a conventional silicon semiconductor detector. Discrete masses of isotopes can clearly be distinguished, which allows their independent projection to the energy axis. Typically, energy signals used in the data analysis are obtained from a timing signal which is converted first to velocity, and after identification of the mass of the isotope in question, to energy.

Energy calibration of a TOF detector is usually very linear, stable, and independent of the detected particle mass. Careful mass calibration [23] allows the plotting of a full mass spectrum from a thin film sample.

In a gas ionisation detector, separation of different atomic species is achieved by detecting differences in their ionisation profiles in gas. This is usually obtained either by collecting the ionisation with an electrical field parallel to the ion tracks and measuring the differences in pulse shapes (Bragg chamber), or by using an electrical field perpendicular to the ion tracks and collecting an electron signal with an anode divided into several parts. The latter type of gas ionisation detector has the great advantage of being position sensitive when total signals of cathode and anode behind a Frisch grid are compared [24]. Thus, a gas detector can easily be have a large solid angle, which makes it very sensitive. A drawback with gas detectors is that their energy calibration is dependent on the atomic species [25]. Also, the thin window necessary between gas volume and chamber vacuum deteriorates energy resolution.

Recently great progress has been made [26] in the use of a combination of TOF and gas ionisation detectors. This kind of device allows the good energy resolution and easy calibration of a timing detector, and the position sensitivity of the gas ionisation detector. It has also became evident that a carefully manufactured gas detector with modern low noise electronics results in much better energy resolution for heavy ions than a silicon solid state detector [26]. This allows good mass separation in a TOF-E detector even at low beam energies. Position sensitivity allows the correcting of kinematic broadening effects, thus yielding a very good depth resolution (below 1 nm) with a relatively simple and low-cost detector setup.

Mass or elemental selective measurement systems require a measurement of more than one parameter in coincidence for each detected particle, i.e., a multi-parameter system with relevant analog electronics. Currently only few such systems are commercially available with easy-to-use software included. Thus, in many laboratories these data acquisition systems are based on home-made devices often dating from an earlier nuclear physics era.

In multiparameter systems data are collected usually in list mode, since this allows a much smaller file size if many parameters are recorded simultaneously with a high resolution (8k is typical). List mode data collection also allows following changes in the sample or measurement conditions during the measurement [27].

Detectors for ion beam measurements are described in more detail in chapter "Ion Spectrometers and Detectors" of this book.

5.2 Data Analysis in Multiparameter HI-ERDA Measurements

Whether a HI-ERDA measurement is performed with a TOF or a gas ionisation detector, the analysis of measurement data starts by selecting and recognizing events belonging to each atomic species or mass. With suitable software the original measurement list file can thus be separated into sublists belonging to each element in the

sample. By comparing the elemental yields in the sublists it is possible to determine if there has been any change in the sample composition during the measurement.

In the next step, calibrations for energy and possible positions or other information about each event are applied. In a favourable case, an analyst now has data where all sample elements are determined and each detected particle has a calibrated energy value. In the case of a measurement from a thin film sample, the average compositions in the film can now be determined directly by comparing the number of detected events for each element and correcting these numbers by the recoiling cross-sections. When depth distributions are to be determined, the situation becomes a bit more complicated, since the information on the particle depth of origin is included in its energy, but the energy depends on the energy loss of primary ions and recoils in the sample, which then depends on the (unknown) sample composition. The analysis, however, can also be performed in this case without prior knowledge of the sample structure with an iterative procedure (see Fig. 5).

The procedure is started by assuming a homogeneous sample composition obtained by correcting the total yields of each element with their corresponding cross-sections. Stopping forces for each recoil ion are then calculated for this sample composition, and the original depth for each detected particle is then calculated using the detected energy, stopping forces, and known measurement geometry. The resulting depth distributions of detected particles are then used to create a more accurate estimate of the sample composition and its stopping conditions. The calculation of the originating depths of particles and the stoppings is then repeated iteratively until the composition does not change significantly. When the sum of the individual elemental depth profiles is scaled to full 100% concentration (assuming that all elements were detected and recognized), it becomes possible to integrate total elemental concentrations over the depth of the thin film. This iterative event-by-event analysis approach allows also including position information in the analysis to correct angular dependent kinematic, cross-section, and path length effects.

Elemental energy spectra obtained from HI-ERDA measurements often show significant broadening caused by multiple and plural scattering of the primary ions and recoiled atoms. For instance, for the structure analysed in Fig. 5, it is difficult to estimate oxygen content in a thin TaN film grown on a thicker SiO_2 layer. To solve this problem, Monte Carlo (MC) simulations can be applied [28, 29]. In MC simulations ion trajectories in the sample are calculated using ion transport code similar to a popular SRIM code [30]. The resulting energy spectra can be compared to experimental spectra and by varying the assumed concentrations in the simulations, the best fit gives a good estimate for the elemental depth distributions (Fig. 6).

6 Analysis Example: High-*k* Materials

The downscaling of complementary metal-oxide-semiconductor (CMOS) devices has led to the situation where SiO_2 thickness has reached its limit, mainly due to gate
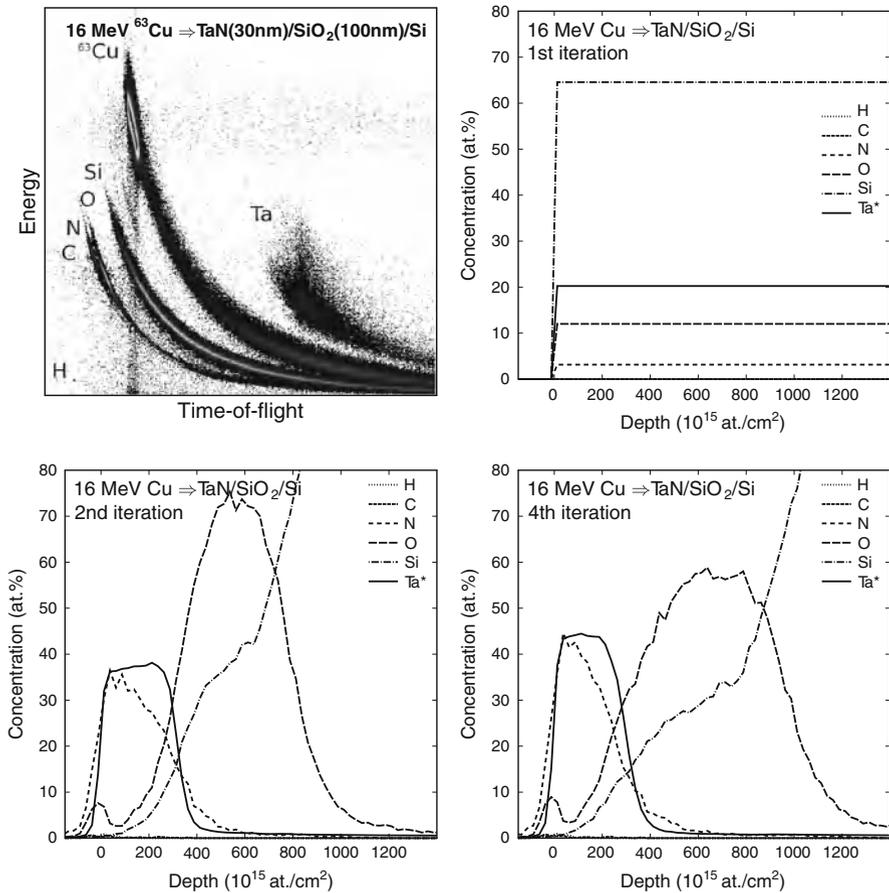
Fig. 5 Iterative procedure for analysis of a HI-ERDA measurement without prior knowledge of the sample structure. See the text for a detailed description of the analysis

leakage current. Because of this, SiO_2 has to be replaced by high di-electric constant (high-k) materials like HfO_2 or $Hf_xSi_{1-x}O_2$. In this search for a replacement for SiO_2, IBA methods have played an important role.

The strongest candidate for the deposition technique, now already in use for chip fabrication by Intel Corporation, is atomic layer deposition. In this technique, precursors carrying the film atoms are subsequently pulsed to the wafer surface, resulting in film growth [1]. The film composition, thickness, impurity contents, and crystallinity are strongly dependent on the precursors, growth temperature, and pulse lengths. Ion beam analysis can reveal the first three of the properties; and, as H and C are many times the key impurities, the information obtained is unique, as shown in Fig. 7. In addition to light impurities, heavy ones like Cl, Br, and I can also originate from the precursors and can be detrimental for the devices, if present.
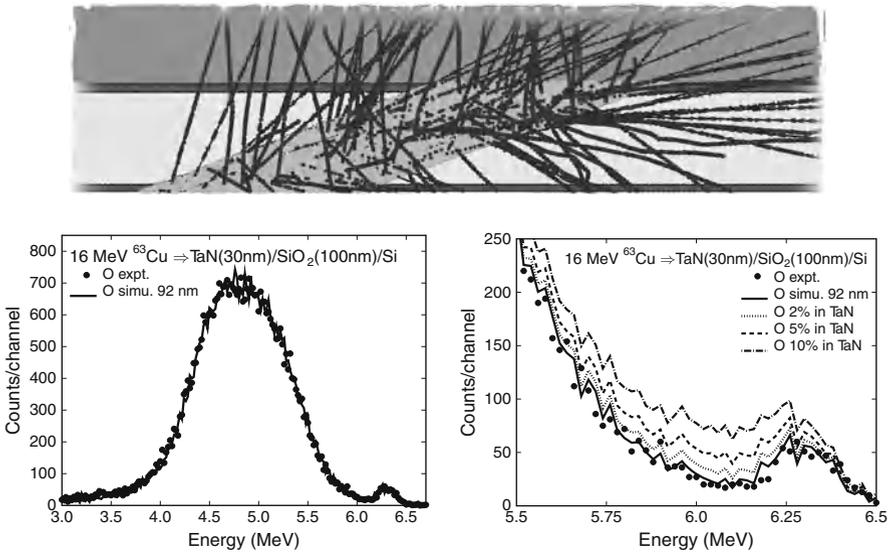
Fig. 6 Use of MC simulations in the analysis of HI-ERDA measurements [28, 29]. Transport of primary ions and recoil atoms are calculated following the particles from collision to collision in the target layers and detector structures (figure above). For the measurement shown in Fig. 5, the oxygen energy spectrum can be simulated in detail by assuming 92 nm SiO_2 layer below oxygen-free TaN film (left). The small high energy peak is caused by surface oxygen contamination, which can also be simulated. By varying assumed oxygen content in TaN film, maximum oxygen content of the film can be estimated (right). In this case it is clearly below 2 at.%
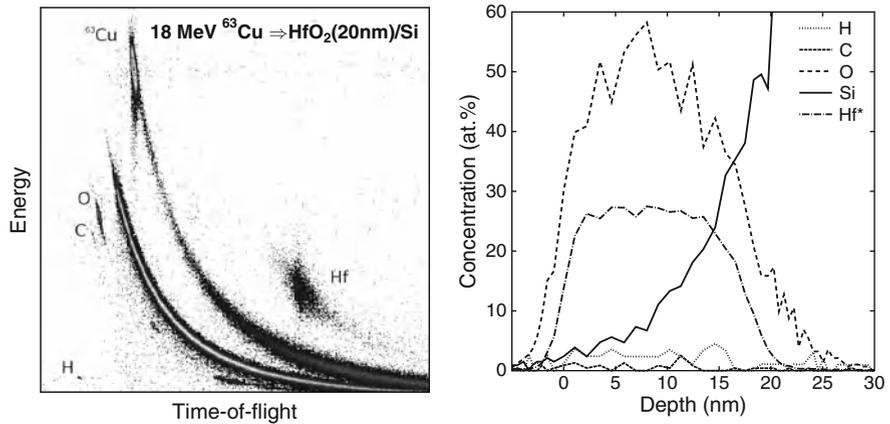
Fig. 7 High-k analysis example. Atomic layer deposited HfO_x film was analysed using TOF-ERDA, and both O/Hf ratio (2.0±0.1) and impurity contents (H 4.4±0.4 at.%, C 1.3±0.3 at.% could be determined

In the depth profiling of films with a thickness of only a few nanometres, and especially with the interface between the grown film and semiconductor, magnetic spectrometers in RBS- [12] and ERDA-mode [19, 31] and medium energy ion scattering (see chapter "Low and Medium Energy Ion Scattering for Near Surface Structure and Nanoscale Depth Profiling") have played a central role. They can reveal the main impurities and their locations and the presence of interfacial oxide.

References

1. M. Ritala, M. Leskelä: Atomic Layer Deposition. In *Handbook of Thin Film Materials* ed. by H.S. Nalwa, Academic Press, San Diego, CA, pp. 103–159, (2002)
2. A. Inouye, S. Yamamoto, S. Nagata, M. Yoshikawa, T. Shikama: Nucl. Instr. Meth. B **266**, 301 (2008)
3. J. L'Ecuyer, C.Brassard, C. Cardinal, J. Chabbal, L. Deschenes, J.P. Labrie, B. Terreault, J.G. Martel, R. Saint Jacques: J. Appl. Phys. **47**, 381 (1976)
4. B.L. Doyle, P.S. Peercy: Appl. Phys. Lett. **34**, 811 (1979)
5. G. Amsel, J.P. Nadai, E. D'Artemare, D. David, E. Girard, J. Moulin: Nucl. Instr. Meth. **92**, 481 (1971)
6. C. Olivier, M. Peisach, T.B. Pierce: J. Radiol. Chem. **32**, 71 (1976)
7. J.-P. Thomas, J. Engerran, A. Charard, J. Tardy: Nucl. Instr. Meth. **119**, 373 (1974)
8. W. Lanford, H. Trautwetter, J. Ziegler, J. Keller: Appl. Phys. Lett. **28**, 566 (1976)
9. P. Torri, J. Keinonen, K. Nordlund: Nucl. Instr. Meth. B **84**, 105 (1994)
10. S.A.E. Johansson, J.L. Campbell: *PIXE: A Novel Technique for Elemental Analysis*, Wiley, New York, (1988)
11. F. Watt, T. Osipowicz, T.F. Choo, I. Orlic, S.M. Tang: Nucl. Instr. Meth. B **136–138**, 313 (1998)
12. K. Kimura, M. Mannami: Nucl. Instr. Meth B. **113**, 270 (1996)
13. W.A. Lanford, S. Bedell, S. Amadon, A. Haberl, W. Skala, B. Hjorvarsson: Nucl. Instr. Meth. B **161–163**, 202 (2000)
14. R. Grötzschel, C. Klein, M. Mäder: Nucl. Instr. Meth. B **219–220**, 344 (2004)
15. M.H. Mendenhall, R.A. Weller: Nucl. Inst. Meth. B **51**, 400 (1990)
16. M. Döbeli, P.C. Haubert, R.P. Livi, S.J. Spicklemire, D.L. Weathers, T.A. Tombrello: Nucl. Instr. Meth. B **56/57**, 764 (1991)
17. K. Kimura, S. Joumori, Y. Oota, K. Nakajima, M. Suzuki: Nucl. Instr. Meth. B **219–220**, 351 (2004)
18. K. Nakajima, Y. Okura, M. Suzuki, K. Kimura: Nucl. Instr. Meth. B **219–220**, 514 (2004)
19. K. Kimura, K. Nakajima, M. Mannami: Nucl. Instr. Meth. B **138–138**, 1196 (1998)
20. Y.S. Kim, J.K. Kim, H.W. Choi, G.D. Kim, H.J. Woo: Nucl. Instr. Meth. B **138-138**, 724 (1998)
21. M. Döbeli, R.M. Ender, V. Liechtenstein, D. Vetterli: Nucl. Instr. Meth. B **142**, 417 (1998)
22. T. Sajavaara, B. Brijs, S. Giangrandi, K. Arstila, A. Vantomme, W. Vandervorst: Nucl. Instr. Meth. B **249**, 292 (2006)
23. M. El Bouanani, H.J. Whitlow, M. Hult, L. Persson, M. Andersson, E. Swietlicki, M. Östling, C. Zaring, P.N. Johnston, S.R. Walker, I.F. Bubb, D.D. Cohen, N. Dytlewski: Nucl. Instr. Meth. B **94**, 530 (1994)
24. W. Assmann, H. Huber, C. Steinhausen, M. Dobler, H. Gluckler, A. Weidinger: Nucl. Instr. Meth. B **89**, 131 (1994)
25. T.D.M. Weijers, T.R. Ophel, H. Timmers, R.G. Elliman: Nucl. Instr. Meth. A **483**, 676 (2002)
26. C. Kottler, M. Döbeli, F. Glaus, M. Suter: Nucl. Instr. Meth. B **248**, 155 (2006)

27. S. Giangrandi, B. Brijs, T. Sajavaara, H. Bender, F. Iacopi, A. Vantomme, W. Vandervorst: Nucl. Instr. Meth. B **249**, 189 (2006)
28. F. Schiettekatte: Nucl. Instr. Meth. B **266**, 1880 (2008)
29. K. Arstila, T. Sajavaara, J. Keinonen: Nucl. Instr. Meth. B **174**, 163 (2001)
30. J.F. Ziegler, J.P. Biersack, U. Littmark: *The Stopping and Range of Ions in Solids*, Pergamon Press, New York, (1984)
31. G. Dollinger, C.M. Frey, A. Bergmaier, T. Faestermann: Nucl. Instr. Meth. B **136–138**, 603 (1998)

Nanoscale Materials Defect Characterisation

Eduardo Alves and Mark Breese

1 Introduction

Nanoscale materials are characterized by the low dimensionality of their building blocks. A large interface area density with an interface area of the order of $1,000 \, cm^2$ is easily found in multiple quantum well structures with $1 \, \mu m$ thickness and a 2–5 nm period. Hence, the interface behaviour will dominate or have a crucial impact on the properties of these structures. Additional defects will be mostly located at such interfaces in the form of planar or surface defects.

Defects have a pronounced effect on all types of material properties. In nanoscale structures, such as single or multiple quantum wells and quantum dots, the presence of a small number of defects could destroy their technological usefulness. Moreover, the defects are of the same length scale as the quantum structures, making structural study a great challenge. Techniques with atomic resolution such as High Resolution Transmission Electron Microscopy (HRTEM) and Atomic Force Microscopy (AFM) allow direct visualisation of the atomic arrangements and could provide such information. Other well-established techniques with many applications to the study of low-dimensional systems are High Resolution X-Ray Diffraction (HRXRD) and Extended X-Ray Absorption Fine Structure (EXAFS). Both techniques use information from the reciprocal space, but only give average values over the volume probed by the photon beam (a large number of nanoparticles or multiple quantum wells are measured). However, neither of these techniques provides information with good depth resolution.

Since structural (crystalline) defects are deviations from the periodic, orderly distribution of atoms in the crystalline lattice, ion channelling analysis is a powerful means of probing these imperfections with depth resolutions in the nanometer range. In addition, the ever-increasing complexity of multiple quantum well structures with periods below one nanometer has lattice mismatches which produce highly strained

E. Alves (✉)
Instituto Tecnológico e Nuclear (ITN), Department of Physics, EN. 10, 2686-953-Sacavém, Portugal
e-mail: ealves@itn.pt

systems. The strain results in a measurable misalignment between the channelling directions of the wells with different lattice parameters. The combination of ion channelling with Rutherford Backscattering Spectrometry gives a combination to measure and correlate strain and composition in compound multiple well structures. In this chapter we focus on the applications of ion channelling in backscattering and transmission modes to study defects and strain in low-dimensional structures and thin layers.

2 Channelling Principles and Defect Dechannelling Models

The channelling effect was first predicted by Stark [1] but it was not until 1963 that Robinson and Oen [2] made the first simulations in which it was observed. Experimental evidence was found soon afterwards [3]. Since then there have been many articles and books in which channelling theory and applications have been discussed extensively ([4–6] and references therein). Channelling occurs when particles are steered by the crystalline potential along the lattice rows or planes in single crystalline samples. To maintain a channelled trajectory, the particles cannot exceed the limiting approach distance, r_{min}, to the strings, which determines the maximum angle of incidence for a stable trajectory. Within the continuum approximation the trajectory of axially channelled particles, Fig. 1a, is governed by the equation [6],

$$E_{\perp} = E\psi^2 + V(r) \tag{1}$$

where $V(r)$ is the continuum potential and $E_{\perp}$ represents the transverse component of the particle energy E and must be conserved to guarantee the same angle before and after the interaction. The maximum value accepted for the incident angle ψ of the particles, in order to achieve a channelled trajectory, is correlated with the
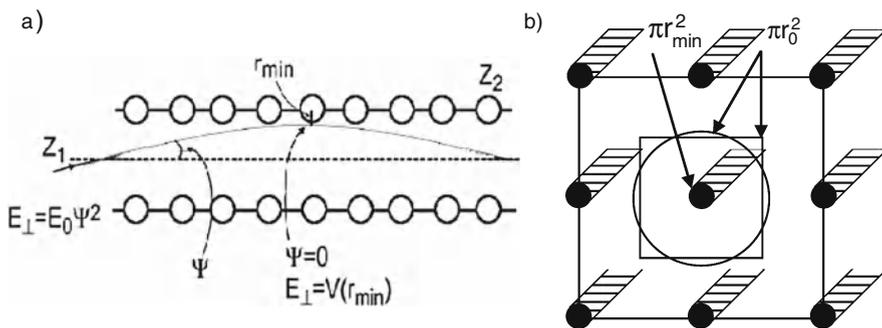
Fig. 1 (a) Schematic of a channelled trajectory incident at the critical angle and having a distance of closest approach, r_{min}. (b) Areas associated with the rows (πr_0^2) and distance of closest approach (πr_{min}^2) whose ratio defines the nonchanelled fraction [5]

closest approach distance, $r_{\min}$, to the rows (Fig. 1a) and is given by [6],

$$\psi_c = (V(r_{\min})/E)^{1/2} \tag{2}$$

The critical angle is sensitive to the displacement of the atoms from their equilibrium positions and is one important quantity in experimental studies. The other relevant parameter is the fraction of dechannelled particles, which is defined as the minimum yield, $\chi_{\min}$, the ratio between the channelled and total beam flux. This parameter is a measure of the close encounter probability, and using simple geometrical considerations in Fig. 1b, is given by [5]

$$\chi_{\min} = Nd\pi r^2 \tag{3}$$

where, N is the atomic density and, d is the atomic spacing along the row. While this simple model indicates a dependence of the minimum yield on material properties, more elaborate treatments [4–6] show a weak dependence on the energy and atomic number of the beam particles.

The potential of ion channelling for defect studies was immediately realized after its discovery and Bøgh [7] set the basis of quantitative defect analysis using the two-beam approximation model (TBA). The model assumes a channelled and a dechannelled beam (the random part of the total beam). According to this approximation, the normalized yield at a depth, x, inside a single crystalline structure is obtained through the expression,

$$\chi(x) = f_R(x) + [1 - f_R(x)]fn_D(x)/N. \tag{4}$$

The atomic density of the material and the density of scattering centres are, N, and, $n_D(x)$, respectively. The parameter, f, is the defect factor with a value between 0 and 1, depending on the type of defects, being almost 0 for dislocations or a perfect crystal and 1 for randomly distributed atoms [5]. The random fraction of the beam, f_R, interacts with all the atoms in the material while the channelled fraction only interacts with the displaced atoms. A knowledge of $f_R(x)$ immediately gives the value of $fn_D(x)$. Several approximations have been used to calculate the dechannelling probability resulting from direct scattering with defects, either using single or multiple scattering processes and considering a defect-dependent cross section, σ_d, for dechannelling [8]. The single scattering theory works well for low concentrations of point defects ($<10^{17}$ cm^{-2}) but its application to identify complex defect structures is rather limited because of the variety of different types of extended or point defects, particularly in nanomaterials normally grown in nonequilibrium conditions where different type of defects may be present. A particular study on misfit dislocations in semiconductor heterostructures was carried out by Mazzer et al. using an analytical model based on the continuum approximation to describe the planar dechannelling [9]. The model describes with good accuracy the experimental results where misfit dislocations were the dominant defect present.

Two other approaches to study dechannelling by point defects or small clusters are the diffusion model [10] and Monte Carlo simulation [11, 12]. Both are widely used in defects studies in irradiated materials and different codes are available. The diffusion model uses Lindhard's theory [13] and solves the equations of motion of the particles considering the displacement of a number of lattice atoms, N_d, into the channels. The density of displaced atoms and the displacement amplitude, r_d, is determined measuring the minimum yield at different temperatures [10].

The Monte Carlo method is based on using the binary collision approximation to calculate the close encounter probability. It is used to simulate the trajectories of the particles and reproduce channelling spectra [12, 14]. Recently Lulli et al. obtained encouraging results with the binary collision approximation using atomic scale models of defects to study implantation damage in Si [15].

While the damage type is rather difficult to identify when various types of defects are present, the defect profile in lightly damaged materials on a nanometer scale is relatively easy to extract. The number of displaced atoms can be obtained using any of the methods described above and the results are well correlated with the microscopic picture of the damaged region, which can be obtained by TEM. An example of the complementarity of the information obtained by channelling and TEM is shown in Fig. 2. The depth distribution of damage produced in GaN implanted with Eu compares well with TEM results, (Fig. 2b) [16, 17]. The nanocrystalline surface contributes to the intense dechannelling peak at the surface (there is always a background contribution from the surface atoms in perfect crystal given by Eq. 3) and the deeply buried damage peak results from the extra dechannelling produced by the stacking faults and atomic clusters.
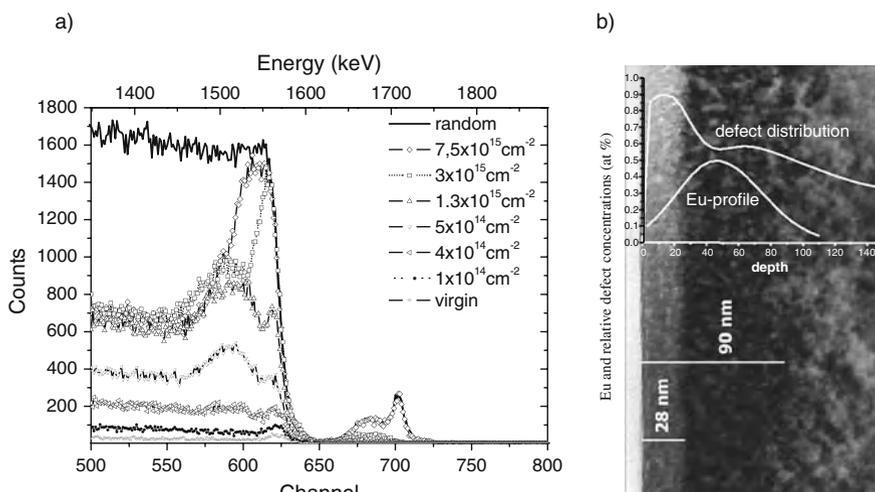
Fig. 2 (**a**) Random and aligned spectra of GaN implanted with different fluences of Eu. (**b**) The damage profiles are well correlated with TEM analysis

3 Strain in Multilayer Systems

The use of RBS to measure the composition independently of the strain state with channelling makes this combination unique in the study of quantum well systems. InGaN/GaN single (SQW) and multiple quantum wells (MQWs) grown by metal-organic chemical vapor deposition (MOCVD) are responsible for the revolution of solid state light sources and the core of a new generation of light emitting devices [18]. Despite all the research activity around this system because of its notable efficiency, the luminescence mechanism is still a controversial issue. The debate is dominated by the existence of phase segregation effects at the nanoscale level [19]. The mismatch between GaN and $In_xGa_{1-x}N$ alloy has a maximum value of 10% for pure InN ($x = 1$). Using RBS/Channelling (RBS-C) to determine the depth profile of the strain in the layers, it is possible to give an interpretation for the double peak observed in XRD and photoluminescence (PL) measurements without the need of compositional inhomogeneity [20]. The well composition, x, was kept constant and the thickness changed to follow the strain release. In these examples, the wells are thick enough to channel the beam particles.

The samples discussed here were grown by MOCVD over $\sim$1.2 μm GaN buffer layers deposited on c-sapphire substrates. The composition determined using the ion Beam Analysis DataFurnace (NDF) code [21] to fit the random spectra, Fig. 3, indicates an In concentration of 9.9 at% and 12 at% for film A and B, respectively. The well thickness of film A is 66 nm and for film B 180 nm. Since RBS measures areal densities (at. cm^{-2}), Vegard's law [22] was used to obtain the well density using the composition measured by RBS.

The strain effect imposed by the lattice mismatch is shown in Fig. 4, where the compression in the basal plane induces an expansion along the c-axis for the strained films. The consequence is a misalignment between the atom rows and planes in both
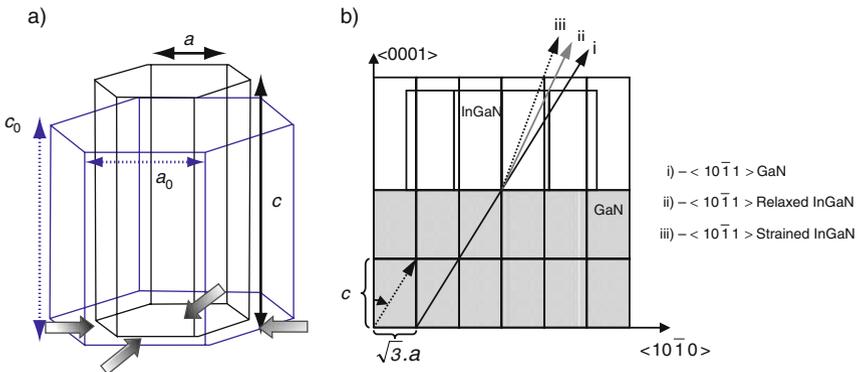
Fig. 3 (**a**) Illustration of the strain effect on epitaxial InGaN layers grown on top of GaN. Because of the fact that the in-plane InGaN lattice constant is larger than that of GaN, the biaxial strain will compress the base of the hexagon and elongate the unit cell in the direction of growth. (**b**) Schematic illustrating the change in tilt in the interface of a $\langle 10\bar{1}1 \rangle$ axis in strained and relaxed InGaN films on GaN

Fig. 4 Random, aligned, and simulated RBS spectra of sample A ($In_{0.1}Ga_{0.9}N/GaN/Al_2O_3(0001)$) and sample B ($In_{0.12}Ga_{0.88}N/GaN/Al_2O_3(0001)$). Vertical arrows indicate the surface scattering energy of the different chemical elements. Horizontal arrow indicates the In depth scale in sample B
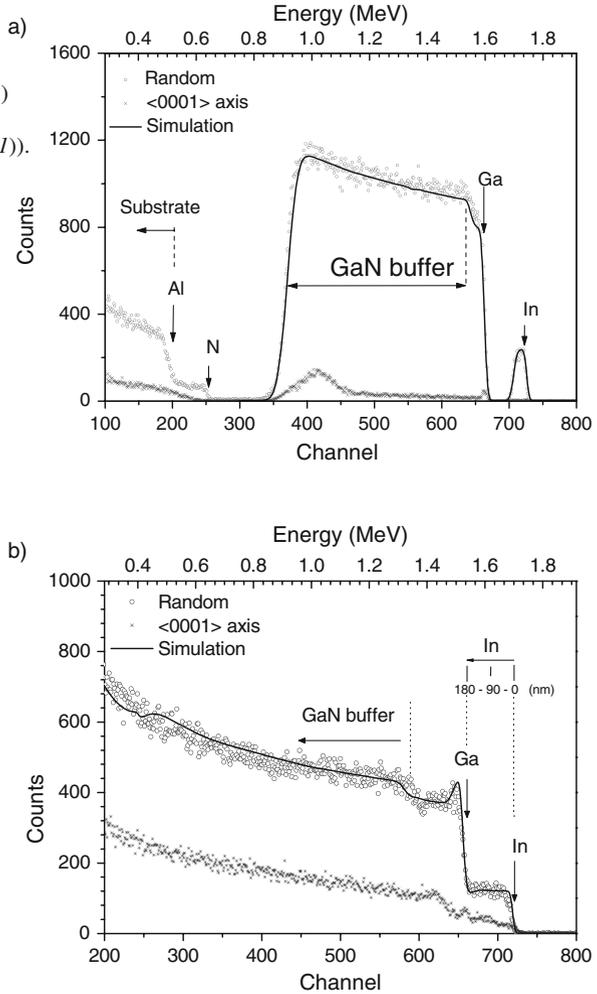
layers, measurable along the tilt directions (Fig. 4). The information on the strain
was obtained through the measurement of the tetragonal distortion [23],

$$\varepsilon_t = \varepsilon_\perp = \varepsilon_\parallel = \frac{\Delta\theta}{\sin\theta\cos\theta} \tag{5}$$

where $\Delta\theta$ is the angular difference between the strained and relaxed film and θ the
tilt angle. The results for samples A and B obtained along the $<10\bar{1}1>$ tilt axis are
shown in Fig. 5. The angular scans were performed along the $\{1\bar{2}10\}$ plane using
the energy windows shown. Sample A exhibits excellent crystalline quality with

Fig. 5 Angular RBS/channelling scans of sample A and sample B, corresponding to the GaN buffer layer and the InGaN layer at the indicated depths, along the $(1\bar{2}10)$ plane from the $<0001>$ to the $<10\bar{1}1>$ crystal axis. The inset shows the angular dependence of the crystal axis on the wurtzite lattice constants
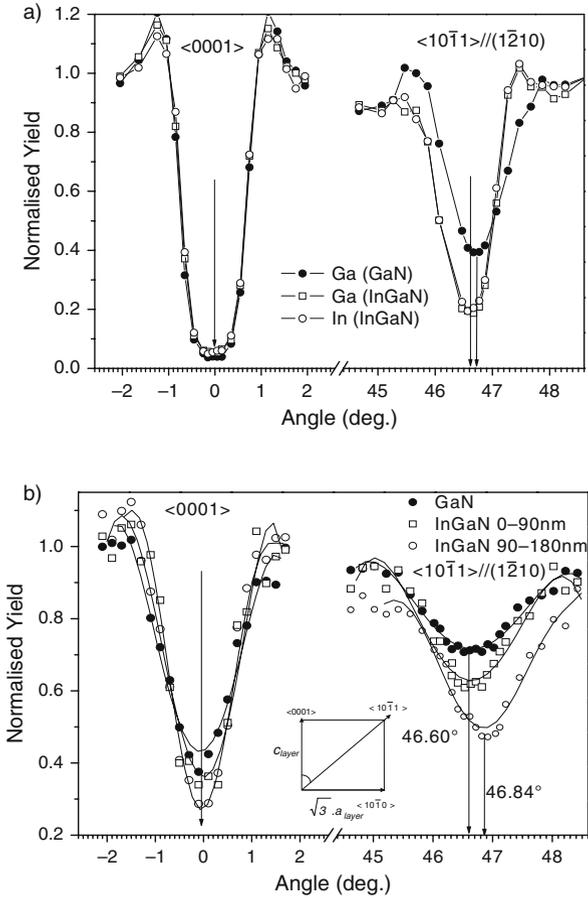
$\chi_{min} \sim 3\%$, comparable to the theoretical value given by Eq. (3) and Monte Carlo simulations. The high minimum yield and lower width of the angular scan along the growth direction indicates a worse quality for sample B. The tetragonal distortion of sample A is -1.04%, while for the two regions defined in sample B (Fig. 5) it is -0.85 and -0.01%.

The minus sign indicates a tensile strain along the c-axis which is higher for sample A. The excellent crystalline quality of sample A measured by channelling means a higher strain, which is less pronounced in sample B, resulting from the presence of defects. A residual value of the tetragonal distortion for the surface region demonstrates a transition from a strain region close to the interface to a relaxed region at the surface. These observations are well-correlated with the asymmetrical reciprocal space maps (RSM) [24] obtained from the XRD data in Fig. 6, where sample B shows two peaks while sample A has only one. PL spectra of sample B also shows two peaks, attributed to phase separation by some authors [25]. Although XRD is

Fig. 6 High resolution XRD reciprocal space map of the (10$\bar{5}$) reflection, for sample B. The vertical line indicates growth coherent to GaN, whereas the dashed line represents the relaxation line. The full line connecting the relaxed, InGaN(R), and strained, InGaN(S), regions is the isocomposition line for $x = 0.19$ [20]
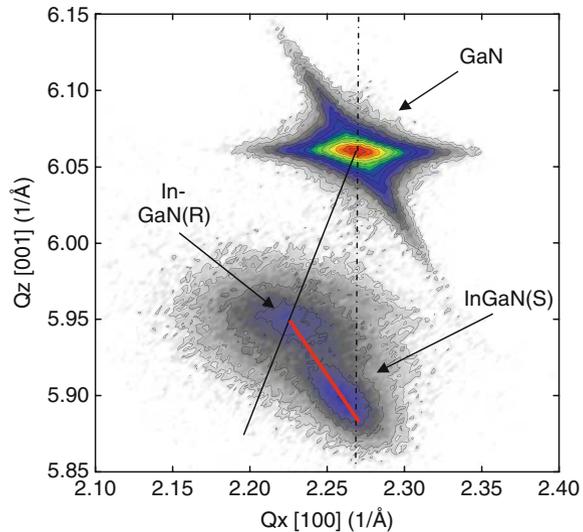
the ideal technique to assess crystalline structure, it relies on the elastic constants and Vegard's law to determine the composition which, associated with the lack of depth resolution, leads to ambiguous interpretation of peaks in the RSM. In fact, the results are explained assuming two different strained regions: a pseudomorphic region highly strained close to the interface and a relaxed region at the surface. Indeed, after removing the surface region by chemical etching, only the peak corresponding to the strained region remains [26]. The results are in good agreement with the critical layer thickness for $In_xGa_{1-x}N$ on GaN estimated using the People and Bean model [27] where relaxation is expected to occur for a thickness above 75 nm for an InN fraction of 10%.

An interesting situation occurs when $\Delta\theta$ is of the same order as the critical angle. In this case the crystalline potential can bend the beam across the interface, maintaining a channelled trajectory, and the measured angular shift does not correspond to the real kink angle. This case was studied for the AlInN/GaN system [28]. Using the Monte Carlo code FLUX [29] to simulate the trajectory of the particles, it was found that only for particle energies above 8 MeV, the steering effect could be neglected, Fig. 7. Simulations for a sample with an InN molar fraction of 15.3% give a value of 0.46° for $\Delta\theta$, while the experimental value is 0.68°, Fig. 8. Using the simulated values an excellent agreement with XRD results was found, and the In concentration ($\sim$17%) where the strain changes from tensile to compressive was determined [28]. Furthermore, a discrepancy in the composition measured by RBS and XRD provides controversy about the validity of Vegard's law for this type of compound and about the insufficient knowledge of material parameters such as the elastic constants and lattice constants used by XRD to determine the composition.

Fig. 7 Monte Carlo
simulations showing the
influence of the steering
effect on the trajectory of the
channeled particles
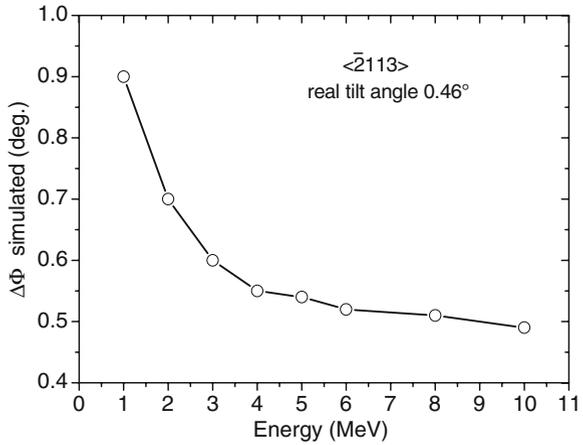
Fig. 8 Experimental angular
scans for Ga and In across the
$\langle \bar{2}113 \rangle$ axis of AlGaN sample
taken with a 2 MeV He$^+$
beam and Monte Carlo
simulations [28]
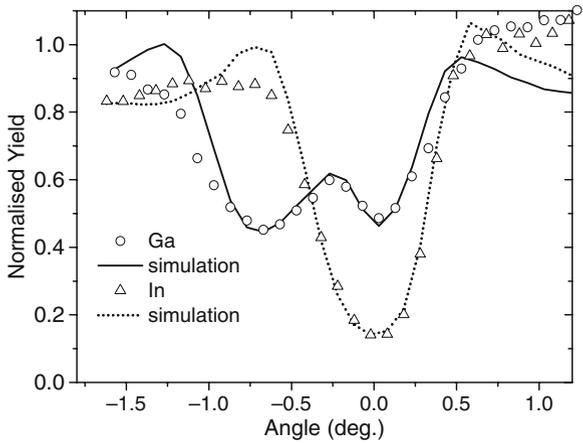
4 Ge Quantum Dots in Si

The growth of Ge quantum dots (QD) in silicon is an attractive route for the fabrication and integration of new optoelectronic devices in well-established Si-based technology. The zero-dimensionality of Ge QD implies 3D confinement of the electrons with electronic states localized and the energy quantized. This has a dramatic effect on the quantum efficiency and thermal stability, making these structures ideal for laser devices and photodetectors. Two growth modes are currently used to grow Ge QDs in silicon, the Stranski-Krastanow (SK) and Volmer-Weber (VW) modes [30]. The SK mode is the characteristic growth mode of strained systems (the lattice mismatch between Si and Ge is 4.16%) where the growth proceeds layer-by-layer during the first atomic layers and then the dots start to form because of the strain accumulation. The formation of 3D islands is a consequence of strain release, and recent studies indicate that it happens for thicknesses above 3 monolayers (ML)

[31]. The VW mode occurs when the lattice mismatch is large enough to inhibit the initial planar growth and the 3D island precipitates. The use of an ultrathin interlayer of SiO_2 before dot formation induces the coherent growth of uniform and small dots (∼10 nm) with a relatively high density (10^{12} at. cm^{-2}) [32]. The lower dimension, resulting in higher confinement, associated with the higher density makes the VW mode very promising. However, while the SK growth mode has been studied for many years, the structural and optical characterization of the dots obtained through the VW mode is just starting [33, 34].

RBS and RBS-C experiments were performed to study the epitaxy and crystalline quality of Ge QD, including their strain state. The potential of ion beam techniques to give information on buried systems of a few monolayers was explored and its

Fig. 9 AFM images of Ge quantum dots grown on 1 monolayer of SiO_2 over silicon. The homogeneous distribution of dots is achieved for the 0.9 nm Ge thick sample

0.3 nm

0.6 nm
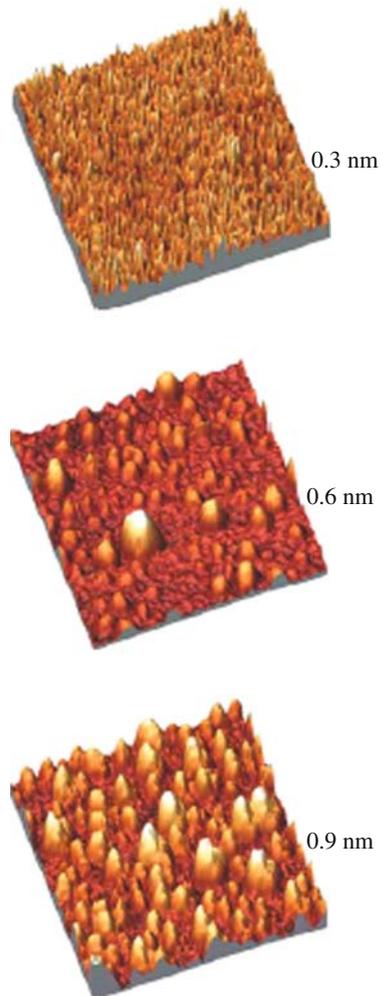
0.9 nm

applicability was proven, where other techniques such as TEM or HRXRD are time consuming or ineffective.

Ge QD were grown by the VW mode with Ge thicknesses of 0.3, 0.6, and 0.9 nm, deposited at $500\,^{\circ}$C on top of 1 ML of SiO_2, Fig. 9. The AFM map of uncapped samples reveals the presence of dots for structures with Ge thicker than 0.6 nm. AFM is unique in allowing visualization of the surface topography with atomic resolution, but does not give any insight on the structural properties. In this particular system, even the use of HRTEM or HRXRD to obtain structural information on deeply buried layers is unrealistic (or even impossible). This structure is ideal for the RBS-C and spectra shown in Fig. 10, which clearly indicates the presence of 0.9 nm of Ge buried under a 90 nm Si cap. The <110> aligned spectrum reveals a peak in the region where the Ge dots and the SiO_2 layer are located. This peak corresponds to displaced Si atoms incorporated in the SiO_2 monolayer and the strain field produced by the dots. No peak was found for the 0.3 nm Ge thickness. The angular scans through the <110> axis along the {100} plane for the 0.3, 0.6, and 0.9 nm Ge thicknesses give evidence of the changes resulting from the formation of the Ge dots, Fig. 11. The overlap between the Si and Ge curves observed for the 0.3 nm sample suggests the incorporation of Ge into substitutional sites of the Si lattice. This conclusion was confirmed by identical behaviour along the other main axial directions. Increasing the Ge content leads to coalescence of the Ge QD with the corresponding perturbation of the crystalline order. PL spectra of these samples showed characteristic emission associated with Ge dots at 0.85 eV [34]. The presence of Ge islands has a pronounced effect on the shape of the Ge angular scans, and a detailed analysis with Monte Carlo simulation is necessary to build a sound interpretation of the data. This is more obvious here since the number of atoms along the Ge rows is insufficient to steer the beam, which is channelled by the potential created by the Si atoms in the cap.

Fig. 10 Random, aligned, and simulated RBS spectra of the structure in the inset. Vertical *arrows* indicate the scattering energy of the different chemical elements at the surface. The presence of displaced Si atoms in the region surrounding the Ge quantum dots (channel 465) is visible in the aligned spectrum
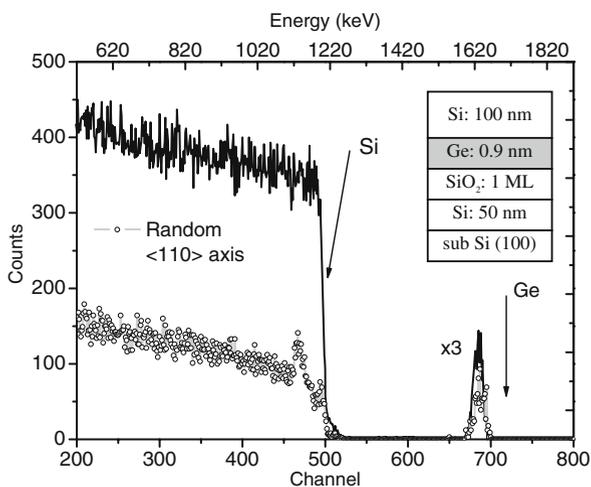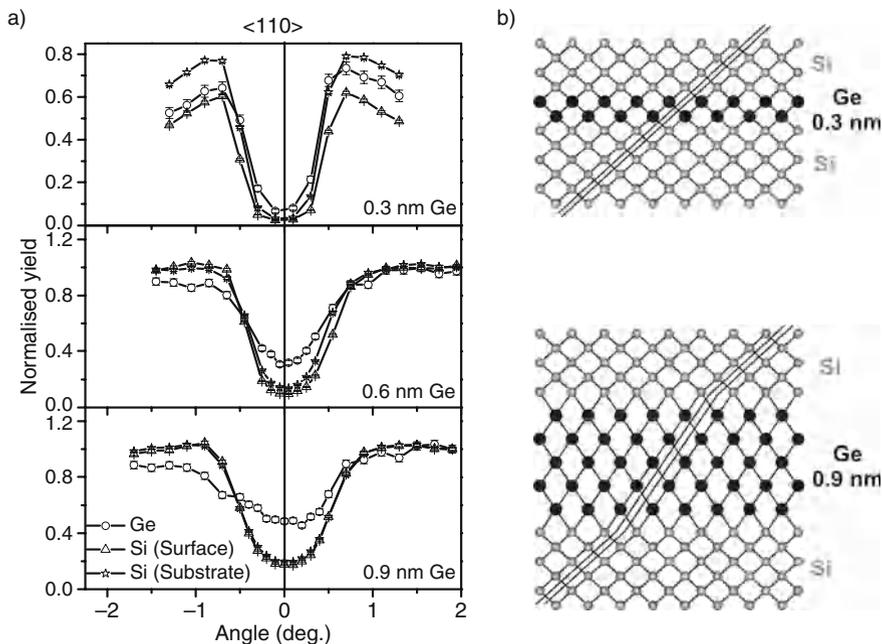
Fig. 11 (**a**) Angular RBS/channelling scans along the <110> axis of Si containing a buried layer of Ge with the thicknesses indicated. (**b**) Atomic diagram for the sample structure showing the effect of Ge atoms on the channeled particles along a tilt direction. The formation of Ge dots could be responsible for the asymmetry of the channeling curve

5 High Resolution Transmission Channelling Analysis

MeV proton beams focused to a spot size of ~ 300 nm in a nuclear microprobe have been used for many years to produce spatially resolved transmission channelling images of defects in thin crystals (see [35] and other references therein). Samples are mechanically thinned and polished to 10–20 μm, i.e., less than the ion range (2 MeV protons have a range of $\sim 50\,\mu$m in silicon). The energy spectrum of the transmitted protons is recorded with a semiconductor detector on the beam axis while the focused beam scans over the crystal surface. Channelled protons lose energy at typically half the rate of nonchannelled protons [36]. Any lattice disruption causes additional dechannelling, and transmission channelling images showing variations in transmitted energy reveal the distribution of defects such as dislocations [37, 38], stacking faults [39, 40], precipitates [41], areas of elastic strain [42], and exfoliated regions [43].

Many effects were not resolved in earlier work because of insufficient spatial resolution and small variations in transmitted energy, resulting in low-contrast, noisy images. These limitations were recently overcome [44] using a new, high-demagnification microprobe which can focus MeV protons to spot sizes as small as 30 nm [45, 46]. Even with large demagnifications the beam convergence angle used
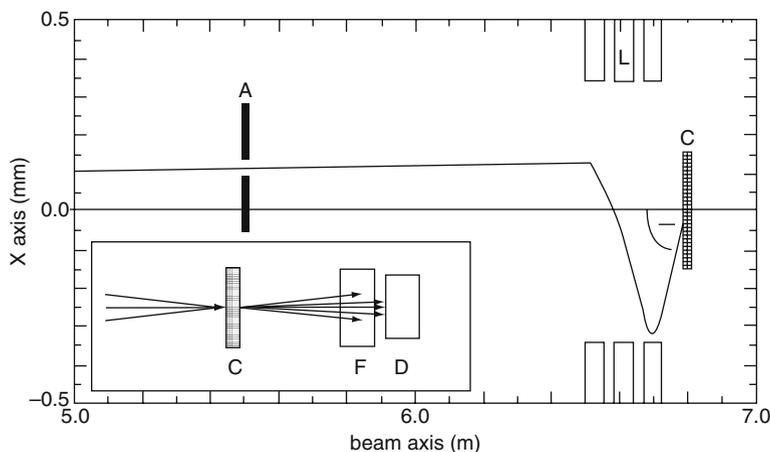
Fig. 12 Trajectory plots for 2 MeV protons focused to a spot on the crystal surface (**C**) by the action of a quadrupole lens triplet (**L**). The direction of negative tilting of the beam with respect to the crystal is shown. When the collimator aperture (**A**) is shifted away from the beam axis, the beam angle to the crystal surface is altered. Inset shows the protons transmitted through the crystal being measured by a particle detector (**D**). The Mylar foil (**F**) ensures that only high-energy, channelled protons are recorded

for transmission channelling measurements is only $\sim 0.01°$, which is much less than $\psi_c \sim 0.17°$ for 2 MeV protons in the silicon (110) planes, because of the very low beam currents of 1 fA to 1 pA needed for analysis.

The image contrast was also greatly increased by placing a foil in front of the detector, as shown in Fig. 12 (inset). The foil thickness was chosen to stop most low-energy, dechannelled protons, so that mainly channelled protons with higher energy and greater range are recorded. Any defect-induced dechannelling now causes significant variations in the measured proton intensity and consequently strong contrast changes in the transmission channelling images.

Altering the Beam/Crystal Alignment

One problem with implementing a transmission channelling facility on a high-demagnification microprobe is the lack of space for a goniometer, owing to the short lens working distance. Beam-rocking systems on nuclear microprobes provide the ability to tilt a focused beam spot through angles of $3°$, with the beam remaining stable on the sample surface to within $2\,\mu m$ [47–49]. However, such systems require careful implementation on a high-demagnification microprobe because the small quadrupole lens bore radius limits the maximum beam deflection angle before entering the lens system.

Another solution is to shift the collimator slits of the microprobe away from the beam axis [50], resulting in the beam being tilted with respect to the crystal

surface, as shown in Fig. 12. For the National University of Singapore (NUS) high-demagnification microprobe, a horizontal/vertical collimator displacement of $50\,\mu m$ produces a change of $0.13°/0.04°$, respectively, in the beam tilt. The location of the beam remains within $1\,\mu m$ and no loss of spatial resolution is observed in a beam spot size of 60 nm for tilts of $0.5°$. This procedure is controllable and can provide very small incremental changes of $0.005°$ in alignment.

6 Transmission Channelling Results

Analysis of Stacking Faults

Figure 13 shows transmission channelling images of a single stacking fault in a [100] silicon wafer at different tilts to the vertically running (011) planes. Stacking faults are 'D-shaped' with the straight edge of each fault meeting the surface along a [011] or [01$\bar{1}$] direction. Dark/bright regions represent regions of high/low energy loss, respectively. The fault appears dark in Fig. 13a because many protons are dechannelled on passing through the fault plane. The additional dechannelling effect of the partial dislocation which bounds the right-hand fault edge is visible as a darker band in the lower-right. In Fig. 13b, the leftmost portion of the fault which is closest to the surface appears as a bright band because of the "blocking to channelling" transition [40]; protons may become channelled at the fault plane because of the lowering of their potential energy. Figure 13c shows a smaller region of the central fault portion at a tilt of $0.10°$ in which periodic bright/dark bands are visible.
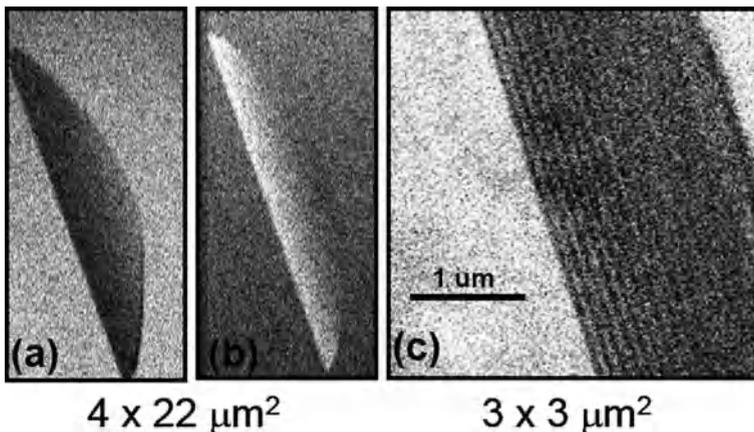
Fig. 13 Transmission channelling images of a stacking fault at tilts of (**a**) $0.00°$, (**b**) $0.22°$, (**c**) $0.10°$. In (**a**) and (**b**) the horizontal × vertical scan area is $4.4 \times 22\,\mu m^2$ and (**c**) is recorded from a $3.3 \times 3.3\,\mu m^2$ area close to the central fault portion. Regions of low/high recorded intensity respectively appear *dark/bright*
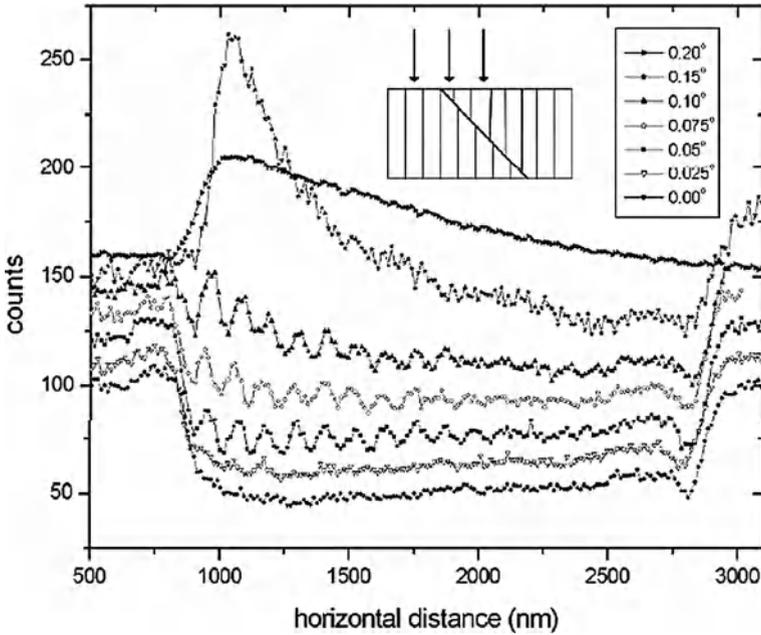
Fig. 14 Horizontal line scans across the centre of the fault in Fig. 13 (**c**) for different beam tilts. In each case the number of counts is normalised to 100 from the virgin crystal at the left-hand edge, and offset by an additional 10 counts for clarity. The inset shows the depth of the fault plane increasing to the right

Figure 14 shows the variation in recorded intensity along a horizontal line scan across the centre of Fig. 13c. For tilts less than 0.05°, the fault region is very smooth and no oscillations are observed. At tilts between 0.05 and 0.10°, periodic intensity oscillations are observed. These were also used to determine the "decoherence length" of the planar channelled protons [44].

Phase Space Model of Planar Channelling

A phase space model of planar channelled ion interactions with stacking faults and lattice rotations was developed [51, 52] that incorporates all observed channelling and blocking phenomena. This model is a more general approach to that described in [53–55] in that it includes both lattice translations and rotations. These can be considered identical for planar channelling, except for a quarter planar oscillation wavelength, $\lambda/4$, in the depth dependence [51].

To explain the oscillations observed across stacking faults, Fig. 15 shows the effect of a lattice translation after the planar channelled beam has passed through a given layer thickness. These simulations with the Monte Carlo code FLUX [29, 56, 57] show the phase space distributions of protons with the perpendicular angle
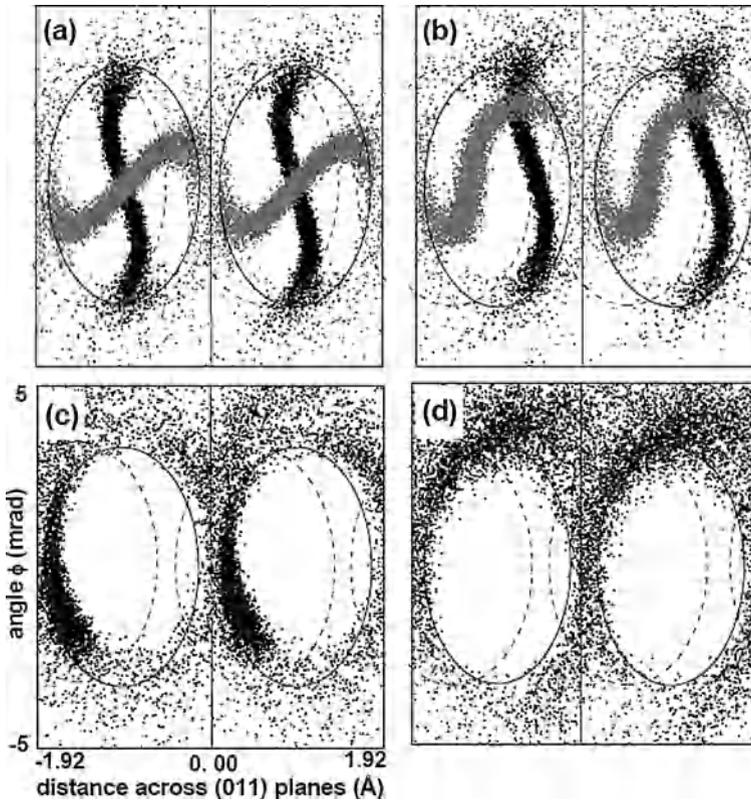
Fig. 15 Phase space plots for beam tilts and layer thicknesses of (**a**) 0.00°, 40 nm (*black* symbols), and 80 nm (*grey* symbols); (**b**) 0.05°, 40 nm (*black*), and 160 nm (*grey*); (**c**) 0.15°, 120 nm; (**d**) 0.20°, 120 nm. The locations of the (011) plane walls and ellipses are shown as *solid lines* within the layer. When the beam encounters the lattice translation after this layer, the displaced ellipses are shown as *dashed lines*

to the (011) planes plotted versus its position with respect to the plane walls. The planar channelled beam occupies an elliptical area which is bounded between the plane walls along the horizontal axis, and by $\psi_c = \pm 3$ mrad along the vertical axis. The lattice translation of the stacking fault after the beam has passed through this layer is incorporated as a horizontal displacement of the ellipses by 0.065 nm.

At planar alignment some channelled protons will be located outside the displaced ellipses, becoming dechannelled. A similar number will be located within adjacent ellipses, producing little net effect, and so no oscillations, as the beam scans over the increasing depth of the fault plane. At small tilts of 0.05°, the phase space distribution is displaced from the ellipse centres and rotates about them. At a layer thickness of 40 nm ($\sim\lambda/4$ where $\lambda \sim 200$ nm for 2 MeV protons) the distribution is close to the right edge of the ellipses; on encountering the lattice translation many protons are located outside the displaced ellipses, giving a minima

in the transmitted intensity. At a layer thickness of 160 nm ($\sim 3\lambda/4$), the distribution is closer to the left edge of the ellipses. On encountering the lattice translation, most protons remain located within the displaced ellipses, giving a maxima in the transmitted intensity and hence oscillations in the line scan in Fig. 14. Similar arguments explain all oscillatory and blocking behaviour observed at stacking faults.

Further details of phase space analysis are given in [51, 52]. Reference [58] describes its use to describe how a lattice rotation at depth intervals of half the oscillation wavelength can result in more low-energy loss trajectories than in a perfect crystal.

Isolated **60°** *Misfit Dislocations*

Transmission channelling images of isolated bunches and individual 60° misfit dislocations in partially-relaxed Si/Si$_{0.95}$Ge$_{0.05}$ layers were recorded [59]. The 60° misfit dislocations are produced in localised regions at the epilayer interface owing to partial relaxation, and run along [110] and [1$\bar{1}$0] directions to form a cross-pattern [60]. The Nomarski image in Fig. 16a shows a cross-arm running along the (110) and (1$\bar{1}$0) planes, and the transmission channelling images in Fig. 16b–d were recorded from within the boxed area. The wide, horizontally-running band produced by the bunch of five dislocations lies along the (110) planes, with two etch pits visible. A change in contrast is observed at this bunch, from dark in Fig. 16b to dark/bright in Fig. 16d.

All components of the Burgers vector of 60° dislocations were considered to fully interpret the results above [60]. A 60° dislocation with a Burgers vector $\mathbf{b} = (a/2)[101]$ and a line direction along [1$\bar{1}$0] can be resolved into edge components $b_1 = (a/4)[110]$, $b_2 = (a/2)[001]$, and a screw component $b_3 = (a/4)[1\bar{1}0]$, where a is the lattice parameter of silicon. The edge component b_1 relieves the interfacial strain, and its primary effect is to distort the (110) planes. If only the larger edge component b_2 is considered, adjacent [1$\bar{1}$0] dislocations in a bunch with the same sign of b_2 comprise a low-angle boundary, resulting in a local rotation of the (110) planes. The lattice plane rotation was calculated as a function of lateral position away from the dislocation core and depth below the epilayer surface because of the combined effects of the two edge components, based on the analytical model described in [9]. The screw component was ignored, since it only distorts the orthogonal (1$\bar{1}$0) planes.

From the behaviour of the lattice plane rotation for different combinations of b_1 and b_2, it was shown that the sign of b_1 determines whether the dislocation appears bright/dark or dark/bright to one side of planar alignment. The sign of b_2 determines at which side of alignment the dislocation will appear dark. Hence all four possible combinations of the two edge components, $\pm b_1$, $\pm b_2$, can be differentiated by noting changes of dislocation contrast on tilting the sample.
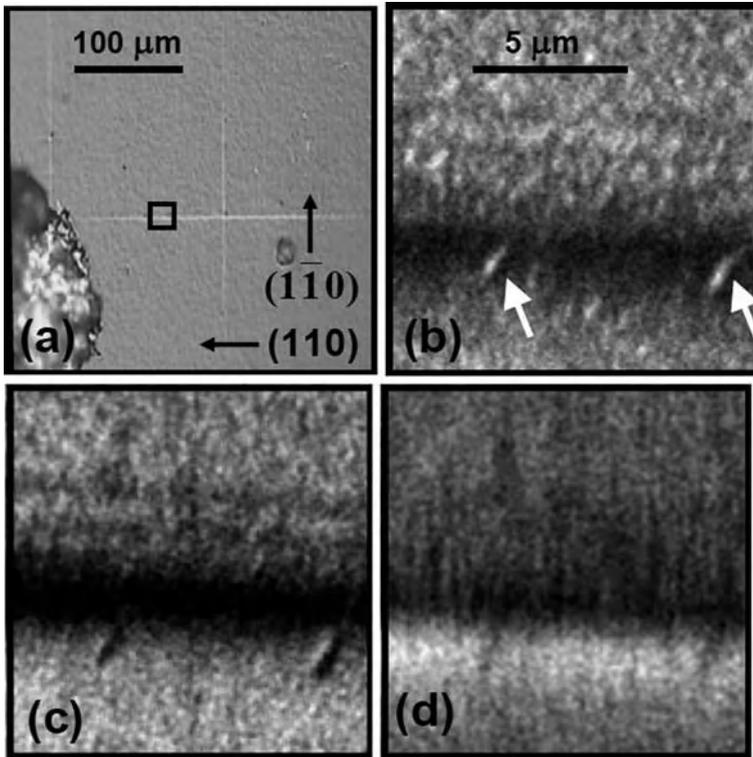
Fig. 16 (**a**) Low magnification Nomarski image. A marker of silver paint was placed at the left edge for subsequent location of this cross. (**b**)–(**d**) Transmission channelling images of the 11 × 11 μm^2 boxed area shown in (**a**). The two arrowed, oval features are etch pits. Beam tilt angle to the (110) planes is (**b**) +0.16°, (**c**) 0.00°, (**d**) −0.16°

7 Summary

In this chapter various applications of channelling to study defects and strain effects have been reviewed. The interpretation of dechannelling to obtain information on defect structures is still a very active research field where new atomistic approaches incorporating relaxation are being considered in recent Monte Carlo simulation codes based on the binary collision model. The broad beam normally used in backscattering experiments is largely compensated by the good depth resolution (1–2 nm at the surface for grazing geometries), allowing the depth profile of defects and strain in complex structures. In transmission mode, channelling reveals the nature and distribution of defect structures inside the materials. With the necessary precautions and a correct description of the interactions, ion beam techniques still have important applications in the study of nanoscale materials because of the importance of their quantitative nature.

References

1. J. Stark, *Phys. Z.*, **13** (1912) 973.
2. M.T. Robinson and O.S. Oen, *Phys. Rev.*, **132** (1963) 2385.
3. R.S. Nelson and M.W. Thompson, *Phil. Mag.*, **8** (1963) 1677.
4. J.A. Davies, Foreign atom location, Channelling: Theory, Observation and Applications, John Wiley & Sons Ltd., New York, Chap. 13, ISBN 0-471-61510-2 (1973) 391–413.
5. L.C. Feldman, J.W. Mayer and S.T. Picraux, Materials Analysis by Ion Channelling: Submicron Crystallography, Academic Press, Inc., ISBN 0-12-252680-5 (1982).
6. D.S. Gemmell, *Rev. Mod. Phys.*, **46** (1974) 129–227.
7. E. Bøgh, *Phys. Rev. Lett.*, **19** (1967) 61.
8. Y. Quéré, *Phys. Status Solidi.*, **30** (1968) 713.
9. M. Mazzer, A.V. Drigo, F. Romanato, G. Salviati and L. Lazzarini, *Phys. Rev.* B, **56** (1997) 6895–6910.
10. K. Gärtner, *Nucl. Instrum. Methods Phys. Res*. B, **132** (1997) 147–158.
11. H. Nakano and Y. Kido, *J. Appl. Phys.*, **71** (1992) 133–139.
12. A. Dygo and A. Turos, *Phys. Rev*. B, **40** (1989) 7704.
13. J. Lindhard, *Dansk. Vid. Selsk. Mat. Fys. Medd.*, **34**(14) (1965).
14. G. Lulli, E. Albertazzi, M. Bianconi, A. Satta, S. Balboni and L. Colombo, *Phys. Rev*. B, **69** (2004) 165216.
15. G. Lulli, E. Albertazzi, M. Bianconi, A. Satta, S. Balboni, L. Colombo and A. Uguzzoni, *Nucl. Instrum. Methods Phys. Res*. B, **230** (2005) 613–618.
16. K. Lorenz, U. Wahl, E. Alves, E. Nogales, S. Dalmasso, R.W. Martin, K.P. O'Donnell, M. Wojdak, A. Braud, T. Monteiro, T. Wojtowicz, P. Ruterana, S. Ruffenach and O. Briot, *Opt. Mat.*, **28** (2006) 750–758.
17. F. Gloux, T. Wojtowicz, P. Ruterana, K. Lorenz and E. Alves, *J. Appl. Phys.*, **100** (2006) 073520.
18. B. Gil (Ed.), Low-Dimensional Nitride Semiconductors, Oxford University Press, Oxford (2002).
19. T.M. Smeeton, M.J. Kappers, J.S. Barnard, M.E. Vickers and C.J. Humphreys, *Appl. Phys. Lett.*, **83** (2003) 5419; A. Rosenauer, D. Gerthsen and V. Potin, *Phys. Status Solidi*. A, **203** (2006) 176.
20. S. Pereira, M.R. Correia, E. Pereira, K.P. O'Donnell, R.W. Martin, M.E. White, E. Alves, A.D. Sequeira and N. Franco, *Mat. Sci. Eng.*, **B93** (2002) 163–167.
21. N.P. Barradas, C. Jeynes and R.P. Webb, *Appl. Phys. Lett.*, **77** (1997) 291.
22. L. Vegard, *Z. Phys.*, **5** (1921) 17.
23. E. Alves, S. Pereira, M.R. Correia, E. Pereira, A.D. Sequeira and N. Franco, *Nucl. Instrum. Methods Phys. Res*. B, **190** (2002) 560–564.
24. S. Pereira, M.R. Correia, E. Pereira, K.P. O'Donnell, E. Alves, A.D. Sequeira, N. Franco, I.M. Watson and C.J. Deatcher, *Appl. Phys. Lett.*, **80** (2002) 3913.
25. S. Pereira, M.R. Correia, E. Alves and K.P. O'Donnell, *Appl. Phys. Lett.*, **87** (2005) 136101.
26. S.M. de Sousa Pereira, K.P. O'Donnell and E. Alves, *Ad. Funct. Mater.*, **17** (2006) 37.
27. R. People and J.C. Bean, *Appl. Phys. Lett.*, **47** (1985) 322.
28. K. Lorentz, N. Franco, E. Alves, I.M. Watson, R.W. Martin and K.P. O'Donnell, *Phys. Rev. Lett.*, **97** (2006) 85501.
29. P.J.M. Smulders and D.O. Boerma, *Nucl. Instrum. Methods Phys. Res*. B, **29** (1987) 471.
30. K.L. Wang, S. Tong, A. Khitun, J.L. Liu and G. Jin, Self-assembled germanium nano-islands on silicon and potential applications, T. Steiner (Ed.) Semiconductor Nanostructures for Optoelectronic Applications, Artech House, Inc., London, Chap. 9.
31. D.J. Eaglesham and M. Cerullo, *Phys. Rev. Lett.*, **64** (1990) 1943.
32. A.A. Shlyaev, M. Shibata and M. Ichikawa, *Phys. Rev*. B, **62** (2000) 1540.
33. A.A. Shklyaev and M. Ichikawa, *Appl. Phys. Lett.*, **80** (2002) 1432.

34. A. Fonseca, E. Alves, J.P. Leitão, M.C. Carmo and A.I. Nikiforov, *Mater. Sci. Eng*. B, **124** (2005) 462.
35. M.B.H. Breese, D.N. Jamieson and P.J.C. King, *Materials Analysis Using a Nuclear Micro-probe*, Wiley, New York (1996).
36. L.C. Feldman and B.R. Appleton, *Phys. Rev. B*, **8** (1973) 935.
37. M.B.H. Breese, P.J.C. King, J. Whitehurst, G.R. Booker, G.W. Grime, F. Watt, L.T. Romano and E.H.C. Parker, *J. Appl. Phys*., **73** (1993) 2640.
38. M.B.H. Breese, P.J.C. King, P.J.M. Smulders and G.W. Grime, *Phys. Rev. B* **51**, (1995) 2742.
39. P.J.C. King, M.B.H. Breese, P.R. Wilshaw and G.W. Grime, *Phys. Rev. B* **51** (1995) 2732.
40. P.J.C. King, M.B.H. Breese, P.J.M. Smulders, P.R. Wilshaw and G.W. Grime, *Phys. Rev. Lett*., **74** (1995) 411.
41. P.J.C. King, M.B.H. Breese, D. Meekeson, P.J.M. Smulders, P.R. Wilshaw and G.W. Grime, *J. Appl. Phys*., **80** (1996) 2671.
42. P.J.C. King, M.B.H. Breese, P.J.M. Smulders, A.J. Wilkinson, G.R. Booker, E.H.C. Parker and G.W. Grime, *Appl. Phys. Lett*., **67** (1995) 3566.
43. M.B.H. Breese, L.C. Alves, T. Hoechbauer and M. Nastasi, *Appl. Phys. Lett*., **77** (2000) 268.
44. M.B.H. Breese, E.J. Teo, M.A. Rana, L. Huang, J.A. van Kan, F. Watt and P.J.C. King, *Phys. Rev. Lett*., **92** (2004) 045503.
45. F. Watt, J.A. van Kan, I. Rajta, A.A. Bettiol, T.F. Choo, M.B.H. Breese and T. Osipowicz, *Nucl. Instrum. Methods B*, **210** (2003) 14.
46. J.A. van Kan, A.A. Bettiol and F. Watt, *Appl. Phys. Lett*., **83** (2003) 1629.
47. D.G. de Kerckhove, M.B.H. Breese and G.W. Grime, *Nucl. Instrum. Methods B*, **129** (1997) 534.
48. D.G. de Kerckhove, M.B.H. Breese, A.G. Wilkinson and G.W. Grime, *Nucl. Instrum. Methods B*, 136–138 (1998) 1240.
49. D.G. de Kerckhove, M.B.H. Breese and G.W. Grime, *Nucl. Instrum. Methods B*, **140** (1998) 199.
50. M.B.H. Breese, P.J.C. King, G.W. Grime, P.J.M. Smulders, L.E. Seiberling and M.A. Boshart, *Phys. Rev. B*, **53** (1996) 8267.
51. M.B.H. Breese and P.J.M. Smulders, *Phys. Rev. Lett*., **81** (1998) 5157.
52. M.B.H. Breese and P.J.M. Smulders, *Nucl. Instrum. Methods B*, **145** (1998) 346.
53. J.A. Ellison, *Nucl, Phys. B*, **206** (1982) 205.
54. S.T. Picraux, W.R. Allen, R.M. Biefeld, J.A. Ellison and W.K. Chu, *Phys. Rev. Lett*., **54** (1985) 2355.
55. W.K. Chu, W.R. Allen, S.T. Picraux, J.A. Ellison, *Phys. Rev. B*, **42** (1990) 5923.
56. P.J.M. Smulders, D.O. Boerma and M. Shaanan, *Nucl. Instrum. Methods B*, **45** (1990) 450.
57. P.J.M. Smulders, *Nucl. Instrum. Methods B*, **94** (1994) 595.
58. M.B.H. Breese, M.A. Rana, T. Osipowicz and E.J. Teo, *Phys. Rev. Lett*., **93** (2004) 105505.
59. M.B.H. Breese, L. Huang, E.J. Teo, P.J.C. King and P.R. Wilshaw, *Appl. Phys. Lett*., **87** (2005) 211907.
60. C.J. Gibbings, C.G. Tuppen and M. Hockly, *Appl. Phys. Lett*., **54** (1989) 148.

Box 6: Nanoscale Defects

Eduardo Alves and Mark Breese

Defects affect virtually all properties of crystalline materials, and their role is magnified in nanoscale structures. In this box we describe the different type of defects with particular emphasis on point and linear defects. Above zero Kelvin all real materials have a defect population within their structure, which affects either their crystalline, electronic or optical properties. It is common to attribute a negative connotation to the presence of defects. However, a perfect silicon crystal or any other defect-free semiconductor would have a limited functionality and might even be useless.

Thermal vibrations of the lattice give rise to an equilibrium concentration of intrinsic defects in solids. The type of defects depends on several factors and can be grouped into two major classes with respect to their thermodynamic nature:

– Reversible defects
– Irreversible defects

Irreversible defects are usually associated with structural changes (twins, mosaicity) or agglomeration of point defects (dislocations, voids, stacking faults).

Reversible defects follow thermodynamic rules, and the energies associated with their formation can be calculated in equilibrium conditions [1, 2]. Whenever a defect is created, there is an increase in lattice energy compensated by a change in the system entropy. At a constant temperature and pressure this process is described by:

$$\Delta G = \Delta H - T\Delta S$$

where ΔG, ΔH, and ΔS are the Gibbs free energy variation, the enthalpy, and entropy changes. A detailed treatment of defect theory is beyond the scope of this book and excellent reviews can be found in [3, 4].

E. Alves (✉)
Department of Physics, Instituto Tecnológico e Nuclear (ITN), EN. 10, 2686-953-Sacavém, Portugal
e-mail: ealves@itn.pt

A good understanding of the physics and chemistry of defects is essential to control their role in materials. This is of paramount importance in nanoscale structures where the active regions of optoelectronic and electronic devices, e.g., transistors and LEDs, are in the monolayer range. Intrinsic point defects can be of the Frenkel type (pairs of vacancies and interstitials of the same chemical species); the Schottky type (vacancies), characteristic of closed packed structures; and antisite defects (species A on site B of an AB compound). The zinc blende structure of GaAs is shown in Fig. 1, where some of these defects are indicated. The interstitial As atom in GaAs, responsible for one of the most challenging metastable defect levels, EL2 [5], was studied with ion channelling [6]. The stable position for interstitial arsenic was determined to be the <111> bond center (BC) site (Fig. 1), and the results are compatible with the site symmetry found by the hyperfine techniques, electron spin resonance (ESR) and electron-nuclear double resonance (ENDOR).

An important issue related to point defects is their charge state. This is particularly important in semiconductors and insulators, where charge compensation processes are necessary to guarantee neutrality. The equilibrium concentration of charged defects and the charge state of defects in semiconductors depends on the Fermi level [7]. This is of great importance in understanding the diffusion processes in semiconductors, where defects are vehicles for atom transport. Diffusion of phosphorus, the n-type dopant *par excellence* in silicon, has strong interactions with silicon self-interstitials and displays long range diffusion tails which have deleterious
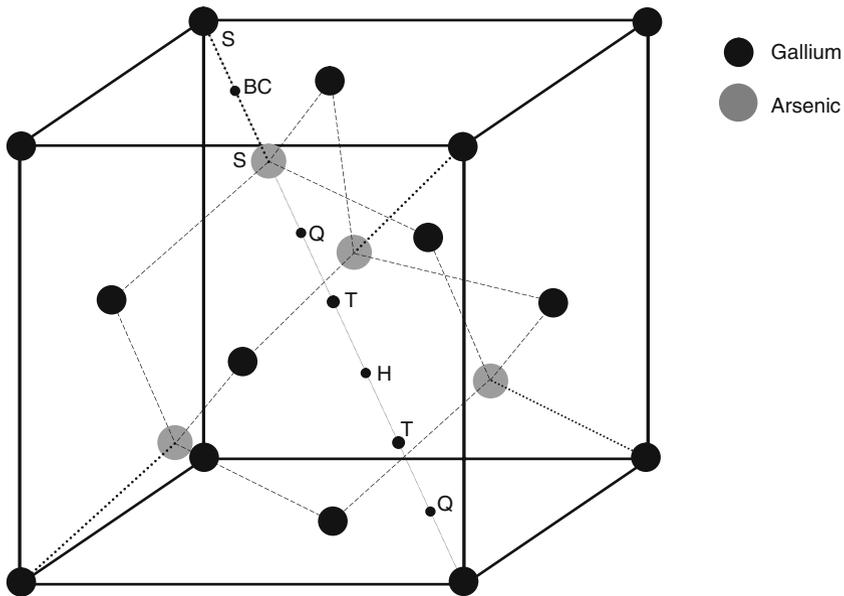
Fig. 1 Conventional unit cell of GaAs lattice, with several impurity sites labeled: substitutional (*S*), bond center (*BC*), antibonding (*Q*), tetrahedral (*T*), hexagonal (*H*), and *C* sites (*C*)

Fig. 2 Projection of the high-symmetry sites in wurtzite GaN lattice on the plane (11–20)
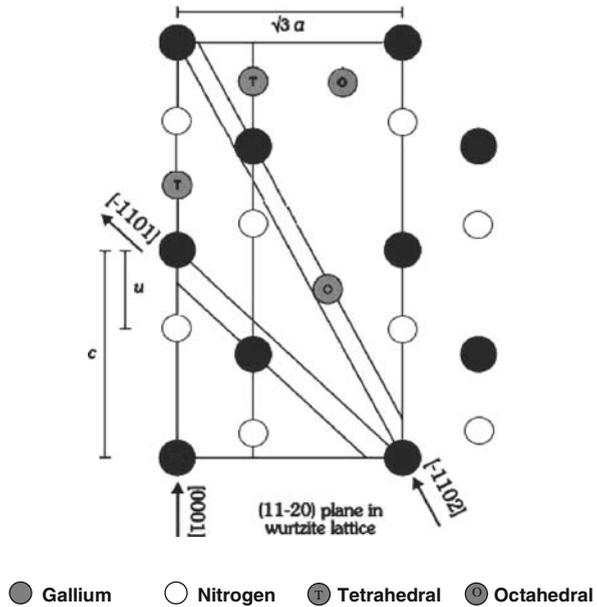
effects on device applications, especially when the dimensions are in the nanoscale range [8].

Besides intentionally introduced impurities, called dopants, there are a vast number of different chemical elements incorporated in material structures during their growth and processing. This important class of defects, known as extrinsic defects, represents the easier route to change the properties of materials in a controlled way. The lattice site location of these impurities determines, in most cases, how the properties of materials are affected. Several high-symmetry sites are normally available to accommodate the impurities in the crystalline lattice, and Fig. 2 shows a projection of these sites for the wurtzite structure of GaN. Since GaN has a natural large n-type background, there were attempts to reduce and control this conductivity. The first attempt to explain the n-type nature of GaN was the presence of N vacancies [9] which, although still subject to some controversy, was recently explained in terms of the presence of impurities such as oxygen and silicon [10]. Many other examples make clear the importance of knowing the site location, thermal stability, and type of impurities, whether they are desirable or unwanted, in order to understand the behavior of materials. The importance of ion beams to provide this kind of information is well documented [11–13]. The use of ion channelling to quantitatively study the site location of any kind of impurity or defect complex is still one the most reliable methods.

It is normal to consider a concentration of 0.1 at.% as the maximum equilibrium value of dilute point defects in materials. For higher concentrations, defects start to interact and complexes are formed. The presence of a large number of

defects is always found in doping processes using ion implantation (see Part IV), or materials under extreme radiation conditions, such as in the walls facing a plasma. The recombination of most of these defects is usually possible by thermal treatment, but in some cases a number of extended defects are formed which are very stable and inhibit complete damage recovery. This is likely in compound semiconductors and insulators with highly ionic bonds. A typical example is GaN where channelling [14] and high resolution transmission electron microscopy (HRTEM) [15] reveal the formation of different types of extended defects during implantation and after annealing.

Different types of extended defects exist, and in most cases they are associated with irreversible processes. Line defects (dislocations) and planar defects (stacking faults, grain boundaries, surfaces, shear planes) are created to release strain fields produced by an excess of point defects or lattice mismatches in systems containing multiple layers of different elements, e.g., superlattices, multiple quantum-wells and other nanostructures.

Dislocations are by far the most studied type of extended defect, and theoretical models have been developed to describe their physical properties, and in particular their movement [16]. The concept of dislocations was introduced in 1934 [17, 18] to
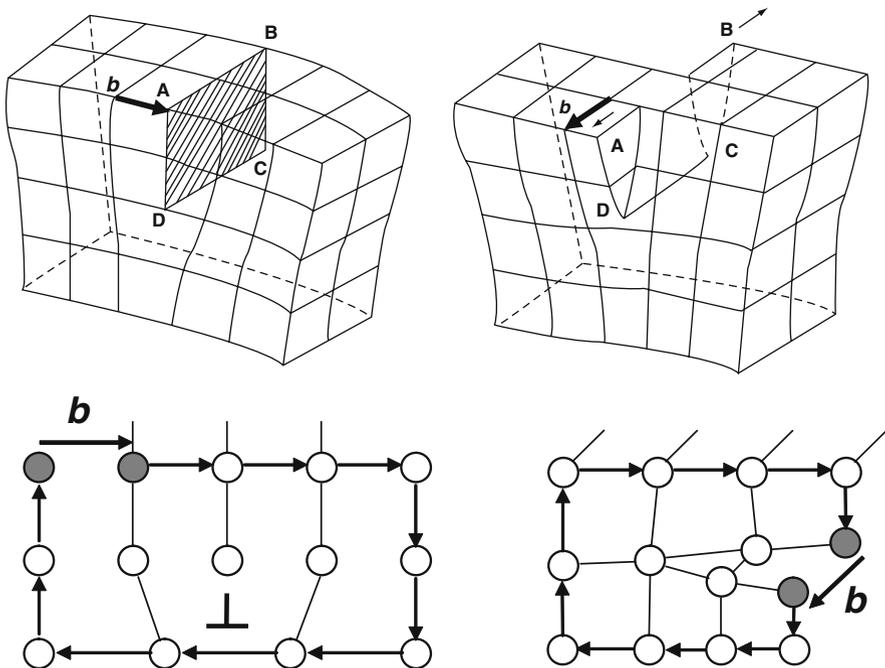
Fig. 3 Scheme of the lattice of a crystal containing **a** edge (*left*) and screw dislocation (*right*); the *arrows* show the direction of the movement of the dislocations and **b** is the Burgers vector

explain the orders of magnitude difference between theoretical and measured values of the strength in crystals. Dislocations are formed during the growth process and are always present in strained multilayer structures above the critical layer thickness. In fact, dislocations are responsible for enhancement of the growth from liquid and gas phases, providing nucleation sites, and for the release of the strain energy accumulated in lattice mismatch systems. In monatomic materials all dislocations are combinations of two extreme types: edge and screw dislocations. Figure 3 shows these two types of dislocations where the characteristic vectors, the Burgers vector, b, and the axis direction, g, are identified. The Burgers vector is perpendicular to the axis vector in the edge dislocations ($b \perp g$), while for screw dislocations it is parallel ($b \| g$). Some other particular angular relations between b and g exist in diamond type lattices where $30°$ and $60°$ dislocations are observed [19]. The presence of dislocations in any type of material and in particular in nanoscale structures, where they can cross different parts of the structure, needs to be considered. The short range ($1/r$) strain field introduced by dislocations attracts point defects and can be useful to precipitate unwanted impurities. The study of this kind of defects with microprobes is discussed in the chapter "Low and Medium Energy Ion Scattering for Near Surface Structure and Nanoscale Depth Profiling."

References

1. S. Mrowec, Defects and Diffusion in Solids: An Introduction, Mater. Sci. Monogr., **5**, Elsevier Science Publ., ISBN 0-444-99776-8 (1980).
2. W. Hayes and A.M. Stoneham, Defects and Defect Processes in Nonmetallic Solids, Lattice Defects, Chap. 3, 105–178.
3. S.M. Hu, Nonequilibrium Point Defects and Diffusion in Silicon, Mater. Sci. Eng., **R13**(3–4), Elsevier Science Inc., ISBN 0927-796X (1994).
4. J.-M. Spaeth, J.R. Niklas and R.H. Bartram, Structural Analysis of Point Defects in Solids (Springer Series in Solid-state Sciences), Springer-Verlag Berlin and Heidelberg GmbH & Co. K, ISBN-10: 3540536159 (1992).
5. J.-I. Nishizawa and Y. Oyama, Stoichiometry of III–V Compounds, Mat. Sci. Eng., **R12**(6–8), Elsevier Science Inc., ISBN 0927-796X (1994).
6. J. Nishizawa, I. Shiota and Y. Oyama, J. Phys. D, **19** (1986) 1073.
7. S.T. Pantelides, Rev. Mod. Phys., **50** (1978) 797–858.
8. S.M. Hue, P. Fahey and R.W. Dutton, J. Appl. Phys., **54** (1983) 6912.
9. Reports on Progress in Physics, Inst. Phys. Publ. **61**(1), ISSN 0034-4885 (1998).
10. J. Neugebauer and C.G. Van de Walle, Appl. Phys. Lett., **69** (1996) 503.
11. J.A. Davies, Channeling: Theory, Observation and Applications, Foreign Atom Location, John Wiley & Sons Ltd., Chap. 13, ISBN 0-471-61510-2 (1973), pp. 391–413.
12. E. Alves, M.F. da Silva, K.R. Evans, C.R. Jones, A.A. Melo and J.C. Soares, Nucl. Instrum. Methods Phys. Res., **B80/81** (1993) 180–183.
13. M.L. Swanson, Handbook of Modern Ion Beam Materials Analysis (eds., J.R. Tesmer and M. Nastasi), Channeling, Chap. 10, University of North Carolina, North Carolina, ISBN 1-55899-254-5, p. 231.
14. S.O. Kucheyev, J.S. Williams and S.J. Pearton, Mat. Sci. Eng., **R33** (2001) 51.
15. F. Gloux, T. Wojtowicz, P. Ruterana, K. Lorenz and E. Alves, J. Appl. Phys. **100** (2006) 073520.

16. R.W.K. Honeycombe, The Plastic Deformation of Metals, 2nd ed., Edward Arnold, London, ISBN 0-7131-3468-2 (1984).
17. G. Taylor, Proc. Roy. Soc., **145** (1934) 362.
18. E. Orowan, Z. Phys., **89** (1934) 660.
19. D. Hull and D.J. Bacon, Introduction to Dislocations, 4th ed., Butterworth-Heinemann, ISBN-10: 0750646810 (2001).

Box 7: Diagnostic Ion Beam Luminescence

Peter Townsend

1 Introduction

Ion implantation of insulating materials generates luminescence, which is routinely used to position an ion beam and, more importantly, provides a powerful diagnostic tool of the host material, both in its initial and in later modified forms. The changes in intensity and emission spectra provide information on the state of the near surface region excited by the ions. The various types of imperfections which exist in the range of the ion beam influence the emission spectra. Signals respond both to original and new defects as well as structures formed during ion implantation. Defect studies are notoriously difficult, as they are always a combination of intrinsic and impurity effects. Current interpretations of defects recognise that many of these features are interlinked. For insulators there is a need to have local charge equilibrium as well as minimizing distortions and lattice stress, and this frequently results in direct clustering of vacancies and impurities. The immediate aggregation of defects into complexes involving tens of lattice sites is not uncommon. Further, there are significant long range electronic interactions, over tens of nearest neighbour shells, so that overall hundreds of ions may be influenced by a complex. All these features can be sensed by luminescence techniques. Hence ion beam luminescence (IBL or IL) is a sensitive monitor of changes in defect structures; and the main difficulty is not in obtaining a signal, but rather in relating it to specific defect sites. One expects that IBL signals from a sample can change in spectral content and intensity as a function of temperature. More surprising is that the signals respond to the presence of amorphous or crystalline phases both of the host material and inclusions such as precipitates or nanoparticles. Early examples of changes in the wavelength of emission from chromium ions in ruby were used to monitor the effective implant temperature during irradiation and implantation [1, 2]. The IBL literature is somewhat sparse and definitely unexploited, which is surprising since luminescence is

P. Townsend (✉)
Science and Technology, University of Sussex, Brighton BN1 9QH, UK
e-mail: p.d.townsend@sussex.ac.uk

simple to detect and analyse and is an inevitable consequence of implantation in insulating materials.

2 Diagnostic Luminescence

Ion beam luminescence has been applied to study defects, impurities, and structural changes in a wide range of materials ranging from crystals to structures used in photonics, as well as natural minerals. The technique has been exploited for studies of glass, thin films, and surface layers modified by ion beam implantation. Signals are also obtained from organic materials, but organics are often very sensitive to radiation damage, which can quench the luminescence.

2.1 Luminescence and Ion Beam Damage

Particularly for light ions (H^+ or He^+), the excitation is predominantly by electronic energy transfer (e.g., see chapter "Basis of Ion Scattering in Nanoscale Materials") and ionisation, hence it resembles cathodoluminescence (CL). Neither CL nor IBL are passive probes of the surface and, depending on the material, either can generate changes in defect structures, colour centres, new luminescence sites, and even total amorphisation of a lattice. The classic example of sensitivity to damage occurs for the alkali halides, where excitonic processes readily form localised displacements on the halogen sublattice. Implantation of organic materials, such as PMMA, will similarly result in major modifications of the target, both from electronic interactions and nuclear collisions. One should note that there has been recent consideration of using ion beams to excite signals from biological materials (see chapter "Nanoscale Engineering in Biosciences"). If the targets are samples of tissue, bone, or teeth, etc., then experimentally the problems are the same as with other samples, but in principle one could use the strong signals associated with in vivo ion beam interactions with surface tissue. The ion beams are inevitably destructive to tissue but in a situation of, say, ion beam removal of skin cancer, then the luminescence signals might offer a valuable and instantaneous marker as to the progress of the treatment.

More stable oxide materials, such as $LiNbO_3$, will primarily only exhibit damage in the nuclear collision region of the implant (i.e., around the end of the range). Nevertheless, electronic excitation causes many crystals and glasses to undergo relaxations, which are immediately apparent from changes in emission spectra. At very high electronic excitation, density defects and particle tracks will develop, even in materials such as $LiNbO_3$. All such changes can lead to more extended defect formation, which also involves the metal sublattice. One notes that for the alkali halides the excitonic damage mechanism is quite temperature sensitive, so it can be partially suppressed by performing the ion beam study at low temperature. Most oxides are more resilient to excitonic damage, by factors of maybe 10^4 for quartz or silica, and even more for alumina and rutile.

2.2 Detection of IBL

Experimentally, studies of ion beam luminescence commenced with visual observations of the signals, which were valuable in beam alignment either when using insulating targets, or by including silica plates on beam-defining apertures. More recently the signals have been used in proton micromachining to visualize the profiled areas of the resist [3]. More quantitative work appeared with the benefit of spectral resolution introduced with a monochromator and photomultiplier (PM) detector. For simplicity these were often bialkali photomultipliers, which have sensitivity over a limited spectral range. More recent systems include high sensitivity PM tubes with spectral coverage from 200 to 900 nm, or CCD systems. The CCD route has the advantages that the detector has a wider spectral coverage to $\sim$1,100 nm and is well suited to inclusion in a wavelength multiplexed spectrometer coupled with a fibre optic to the implantation chamber. The multiplex feature offsets the lower sensitivity of a CCD compared with a PM tube at short wavelengths. In principle, IBL is valuable for longer wavelength analyses of narrow bandgap semiconductors, but currently the detector efficiencies are still poor compared with those in the UV/visible spectral region. However, with development of improved detectors, a vast new regime of experiments and applications would be feasible.

2.3 Comparison Between IBL and Cathodoluminescence

Since cathodoluminescence (CL) analyses are widely used, it is instructive to note that IBL offers numerous advantages compared with CL. These benefits include:

(i) Depth profiling as a function of energy offers a more uniform and controllable rate of electronic energy deposition (dE/dx_{el}) than is possible with CL.

(ii) The dE/dx_{el} value can be varied with a choice of energy and ion species. Of particular value is that comparisons can be made between ion and molecular beam irradiation. There are very large differences in excitation rate, since the molecules dissociate at the surface and the excitation tracks of the fragments overlap both spatially and temporally. Adjusting the energies and particle currents of the ions and molecules allows one to maintain the same total power input and the same projected range. Such a situation immediately identifies dose rate-dependent responses. The high excitation values are unachievable in CL or high current IBL with single ion beams. Comparisons of H^+ and H_2^+, or N^+ and N_2^+, excitation has discriminated between luminescence emission and absorption bands from isolated and larger defect species. In the earlier work at Sussex [4], examples included the ability to alter the relative concentrations of F and F_2 colour centres, as well as H and V_k sites in alkali halides, and F and F^+ defects in alumina, etc.

(iii) With H^+ ions of up to $\sim$3 MeV one can typically measure depth dependent changes to some 10 μm below the surface.

(iv) There are changes in the emission spectra during implantation which directly correlate with formation of defect structures, phase changes, precipitate formation, or amorphisation of a crystalline lattice. During dynamic measurements with continuous recording during heating or cooling, there can be sharply defined changes in spectra, or intensity, which mark the onset of phase transitions [5]. Such events can sometimes be stimulated by the ion beam excitation because one is probing a near surface region, where phase relaxation is particularly readily stimulated by ion beam damage and excitation.

(v) There can be depth dependent variations in luminescence signals in terms of intensity and wavelength. These sensitive responses offer a way to monitor changes in ion implanted optical waveguides, waveguide lasers [4–7], size distributions of nanoparticles [3], and interdiffusion of ions (as in feldspar materials) [8, 9]. Rare earth ions, which are already present in the host, or are introduced as probes, often have line spectra sensitive to local crystal fields. They are therefore ideal, as they tend to be efficient sources of luminescence and so are particularly sensitive when monitoring structural relaxations or other changes. For example, in Nd:YAG waveguides formed by ion beam implants of He^+, one can see changes in the Nd emission wavelengths and relative line intensities. Because of this it is possible to select different lasing lines. Stimulation with a range of energies gives information on pressure changes with depth, nonuniform ingress of water, or other contaminants, and surface dislocations.

References to articles which have employed these concepts span more than 40 years and many early items are discussed in the book by Townsend, Chandler, and Zhang [4] or more recent articles [6–13].

2.4 IBL of Minerals

Data on IBL of minerals are quite limited, but often the challenges are greater than for synthetic crystals because there may be many inclusions of regions of varying composition and, inevitably, minerals are less pure than the laboratory grown materials. One can exploit the impurity variations to image and identify different zones or grains, and such studies have often been combined with other analysis techniques such as PIXE or RBS. Earlier reviews cite examples which exploit the characteristic line emission features of rare earth ion impurities in the minerals [14, 15]. The work from recent years has also shown that one can highlight the presence of trace impurities, such as rare earth ions, and separate changes in the luminescence spectra from sample nonuniformity. Using ion beam luminescence over a range of temperatures, and/or varying the ion species, has enabled the detection of diverse variants of the same basic defect types which occur as the result of impure or stressed compositions. Some of these changes are definitely depth sensitive. In materials such as feldspars there are intrusions and mixtures of different subphases of the family of feldspar compositions. Implantation can record these differences, together with the effects of radiation-enhanced diffusion which both vary

the compositions of the subphases and lead to phase mixing. Thus IBL data offer insights into those processes occurring during mineralogical growth. One recent mineralogical example has included the measurement of IBL for "pure" albite with different growth habits [16]. Here the component emission bands are basically the same, but their relative intensities and temperature responses are quite different. One assumes the IBL is thus responding to stress gradients which occur during crystal growth.

2.5 Phase Transitions and Luminescence

Phase transitions occur in many insulators, and are extremely common in the crystal structures which offer interesting photonic crystal properties, such as nonlinearity or electro-optic responses. Conventionally, such phase transitions have been sensed by crystallographic data. Alternatively, a rapid survey method can be used which records luminescence signals as a function of temperature during a cooling or heating cycle. Excellent examples have emerged from the Sussex group for a wide range of insulators using RL (radioluminescence with X-rays), CL (cathodoluminescence), and even TL (thermoluminescence), in which the spectra, intensity, and luminescence lifetimes all undergo discreet changes at the phase transition temperature. Contemporary Physics A recent review [5] presents numerous examples of phase transitions detected by luminescence, but IBL examples are still lacking or have not yet been attempted.

Examples include evidence for hysteresis effects on the scale of $20°$ for, say, $KNbO_3$. There are also extreme examples where reversal of low temperature transitions requires heating to a high temperature of some $800\,°C$, as for KTP and RTP ($KTiOPO_4$ and $RbTiOPO_4$). Initially such data were monitored for synthetic crystals, but then the method was extended to use implantation to both monitor and induce phase transitions or structural relaxations. In some cases, as for the feldspars, the processes will vary with the host temperature, the rate of energy deposition, and defect formation, but in other cases the transitions will be characteristic of the host lattice and/or impurity phase inclusions. The latter can have spectacular effects on the luminescence intensity of the host (as well as features from impurity phases).

Phase transitions are particularly readily evidenced with CL, since the surface is unconfined and able to relax into alternative structures, even if the transitions are difficult to induce in the bulk. One assumes that with ion beam excitation, the high rates of energy transfer, combined with defects formed by nuclear collisions, will provide catalysts in terms of the local stresses and distortions that trigger the transitions. This opportunity to use ion beams to both induce structural changes and to sense them by luminescence offers a wide range of novel experiments which have not yet received publication in the open literature. More surprisingly, experiments with strontium titanate showed that implantation of high ion fluences into the surface of the crystal could transform not only the CL [17], but also the bulk luminescence responses of RL [18] and TL [19]. The model proposed was that, for the slightly unstable structures into which $SrTiO_3$ can relax, the stress fields

trigger a bulk relaxation. More recently, a similar pattern has been observed for ZnO implants; and one assumes that many other materials, such as silicon carbide, potassium tantalum niobate, or bismuth germinate, could undergo similar behaviour, since they all have multiple phases, variations in their stoichiometry, and high defect densities. For $SrTiO_3$ the evidence was by RL or CL post implantation, but one could readily include simultaneous measurements of RL or photoluminescence (PL) during bombardment. Inclusions of small concentrations of rare earth ions, implanted in the opposite (rear) face from the side receiving the high ion beam fluence, could also provide an optical probe of the structural relaxation of the underlying bulk lattice, by sensing the variations induced in the PL spectra as the front face is bombarded.

Discussions of phase transitions should not be limited to those of the host material, since many impurities segregate from the host and can be viewed as nanoparticle inclusions. These in turn may undergo phase transitions with changing temperature. Associated with changes of the inclusions will be stress fields and long range interactions, often over tens of nearest neighbour distances [20], which dramatically modify properties of the host lattice. In particular these effects are extremely well determined via the RL and CL luminescence signals from the host material. Examples include 100-fold intensity changes of Nd in Nd:YAG which correspond to the sublimation of nanoparticle inclusions of CO_2 near 200 K [21]. Many examples exist that evidence nanoparticle size inclusions of water [22], with most examples resulting in luminescence intensity steps near 170 K (for the cubic to hexagonal phase transition of ice), and these are often accompanied by a further feature near 100 K. Similar evidence for inclusions and phase transitions of oxygen and nitrogen have also been reported, as well as for Ar trapped from a background gas used in pulsed laser deposition of thin films [23]. Equivalent features are to be expected in temperature-dependent measurements of IBL, but so far no examples have been published.

2.6 Future Prospects

Basically the diagnostic aspects of the IBL field have scarcely been developed, although sufficient examples have been realised to show that there are valuable and fruitful opportunities. However, one should note that simultaneous excitation by, say, IBL and/or RL or PL does not necessarily result in the sum of luminescence features from the two independent processes. Our unpublished data show evidence for both additional emission bands and shifts in their temperature dependence compared with independent excitation methods. One assumes that since different defect structures are retained as a function of implant temperature, one can correlate these features by measurements of the IBL emission changes resulting from variations in complex defects and clusters, as well as from RL or PL excitation of transient defects which are formed during the implantation.

3 Conclusions

Ion beam luminescence is a highly informative tool for analysis of impurities and defect structures. Once this is more widely recognised, it will become a standard technique for research on ion implantation in insulating materials. In principle, monitoring of longer wavelength signals will enable it to be applied to narrower band gap semiconductors. There is also the possibility of IBL in biomedical applications which, if effective, would represent an entire new vista for the technique.

References

1. Mitchell EWJ and Townsend PD, Proc Phys Soc, **81**, 12–14, 1963.
2. Chandler PJ and Townsend PD, Rad Effects Letts, **43**, 61–64, 1979.
3. Udalagama CNB, Bettiol AA, van Kan JA and Watt F, Nucl Instrum Methods B, **210**, 256–259, 2003.
4. Townsend PD, Chandler PJ and Zhang L, Optical effects of ion implantation, Cambridge University Press, Cambridge, 1994 (reprinted 2006).
5. Townsend PD, Yang B and Wang Y, Contemporary Phys, **49**, 255–280, 2008.
6. Herreros B, Lifante G, Cusso F, Townsend PD and Chandler PJ, J Phys D, **28**, 1687–1692, 1995.
7. Vázquez G, Rickards J, Lifante G, Domenech M and Cantelar E, Opt Express, **11**, 1291–1296, 2003.
8. Townsend PD, Brooks R, Hole DE, Wu Z, Turkler A, Can N, Suarez-Garcia A and Gonzalo J, Appl Phys B, **73**, 345–353, 2001.
9. Brooks RJ, Finch AA, Hole DE, Townsend PD and Wu Z, Contrib Mineral. Petrol., **143**, 484–494, 2002.
10. Herreros B, Lifante G, Cusso F, Kling A, Soares JC, da Silva MF, Townsend PD and Chandler PJ, J Phys Condens Matter, **10**, 3275–3283, 1998.
11. Brooks RJ, Hole DE and Townsend PD, Nucl Instrum Methods B, **190**, 136–140, 2002.
12. Wu Z, Turkler A, Brooks R, Hole DE, Townsend PD, Koster SF, Kurt K, Gonzalo J and Suarez-Garcia A, Nucl Instrum Methods B, **191**, 121–126, 2002.
13. Finch AA, Garcia-Guinea J, Hole DE, Townsend PD and Hanchar JM, J Phys D, **37**, 2795–2803, 2004.
14. Utui RJ, Homman NPO, Yang C, Malmqvist KG and Tembe SS, Nucl Instrum Methods B, **104**, 432–436, 1995.
15. Pallon J, Yang C, Utui RJ, Elfman M, Malmqvist KG, Kristiansson P and Sjöland KA, Nucl Instrum Methods B, **130**, 199–203, 1997.
16. Garcia-Guinea J, Finch AA, Can N, Hole DE and Townsend PD, Phys Stat Sol (c), **4**, 910–913, 2007.
17. Yang B, Townsend PD and Fromknecht R, J Phys Condens Matter, **16**, 8377–8386, 2004.
18. Yang, B, Townsend PD and Fromknecht R, Nucl Instrum Methods B, **217**, 60–64, 2004.
19. Yang B, Townsend PD, Fan Y and Fromknecht R, Nucl Instrum Methods B, **226**, 549–555, 2004.
20. Townsend PD, Jazmati AK, Karali T, Maghrabi M, Raymond SG and Yang B, J Phys Condens Matter, **13**, 2211–2224, 2001.
21. Maghrabi M, Townsend PD and Vazquez G, J Phys Condens Matter, **13**, 2497–2515, 2001.
22. Kurt K, Ramachandran V, Maghrabi M, Townsend PD and Yang B, J Phys Condens Matter, **14**, 4319–4328, 2002.
23. Brooks RJ, Hole DE, Townsend PD, Wu Z, Gonzalo J, Suarez-Garcia A and Knott P, Nucl Instrum Methods B, **190**, 709–713, 2002.

Nanomaterials Science with Radioactive Ion Beams

Jyrki Räisänen

1 Introduction

Because of the size and mass of nanostructures, their electronic and magnetic property studies with conventional bulk measurement techniques are difficult or even impossible. New techniques, specifically designed for such tasks, are required. The use of radioactive ion beams for studying material properties is not yet widespread but is increasing. An important advantage of employing radioactive nuclei is the ability of detecting signals from a very low concentration of atomic impurities. It is evident that this has essential importance for the characterization, e.g., of semiconductors, because very low amounts of impurities may alter the system properties significantly. Naturally the radioactive isotopes used as dopant atoms also influence the electronic and optical properties of semiconductors according to their chemical nature. The radioactive probe atoms provide us information on the local microscopic environment via observation of the change in the decay characteristics of the unstable species. The potential of employing radioactive ion beam–based techniques for nanoscience applications has not been fully exploited to date. The use of radioactive ions in solid state research, however, dates back to the early 1960s. In pioneering works at the Chalk River Nuclear Laboratories, Canada, the effect of crystalline structure on range distributions were studied, employing 40 keV ^{85}Kr ion implantation [1].

This chapter provides an introduction to the techniques for generating beams of unstable nuclei, introduces the most common techniques based on radioactive ion beams used for probing materials at the nano- and even atomic scale, and solving problems in materials science. Recent examples of the novel uses of the techniques are provided.

J. Räisänen (✉)
Department of Physics, University of Helsinki, FIN-00014, Finland
e-mail: jyrki.raisanen@helsinki.fi

2 Generation of Radioactive Ion Beams

The radioactive isotopes required can be produced at reactors or using particle accelerators. Laboratories having long traditions in employing the former procedure with off-line implanters are the Helmholtz Institut für Strahlen- und Kernphysik at Bonn University and the II Institute of Physics, University of Göttingen, Germany. In this presentation the latter procedure will be looked into in more detail and an introduction is provided to the main techniques employing particle accelerators for generating beams of unstable nuclei, the *ISOL-Isotope Separation On Line* and *In-Flight Techniques*. In all procedures, long-lived isotopes are produced by ion bombardment of target material followed by radiochemical separation from the target material. The isotopes produced can then be either diffused or implanted into the material under study. The most versatile and common procedure is ion implantation.

2.1 ISOL-Isotope Separation On Line

The ISOL method is based on the production of radioactive nuclei via target fragmentation reactions induced by a high-energy light particle beam. The nuclear reaction products produced in the target are extracted as ions and are first stopped, then accelerated (typically to an energy of few tens of keV) and mass-separated in a magnet of an isotope separator on line. The various ISOL concepts differ in the way the reaction products are stopped. The target can be much thicker than the recoil range of the reaction products and most of the products are stopped in the target. In this case the release of the radioactive nuclei relies on diffusion and effusion in the target, which is maintained at high temperatures ($\sim 2500\,^{\circ}$C) to speedup the process. Such diffusion processes vary much in speed. For short-lived nuclear species, with half-lives of milliseconds or less, this is often the intensity limiting factor, because the atoms decay before they reach the final target. Also refractory elements such as zirconium and molybdenum are extremely difficult to ionise and are not available using this procedure at present. In an alternative approach using thin targets, most reaction products recoil out of the target and are stopped in a solid or gaseous catcher which often is an integral part of the ion source. A frequently employed method is to stop the recoil products in high-pressure helium and transport them through a capillary to the ion source. In the so-called ion guide method the gas flows through a differential pumping system is simply directed straight into the acceleration stage of the mass separator and no actual ion source is needed if the nuclei are still in an ionized state. The ISOL method used for producing radioactive ion beams is shown in Fig. 1. Current developments are striving towards more selective systems and targets capable of coping with very high power. A review on the thin target ISOL method may be found from [3].
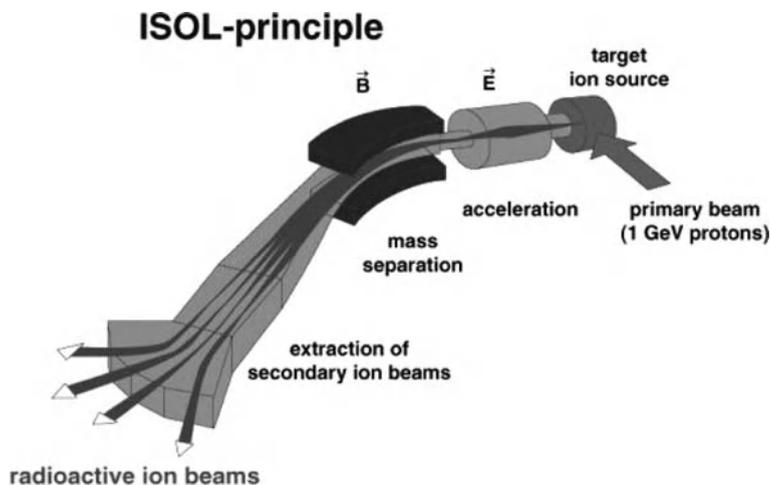
Fig. 1 Schematic view of the ISOL method at Cern/ISOLDE. Radioactive nuclides are produced by spallation, fission, or fragmentation reactions triggered within a thick target by an external proton beam of 1 GeV. The target is kept at high temperatures, permitting the rapid diffusion of the reaction products into the ion source where they are ionized by plasma, laser excitation, or surface ionization. The singly charged ions are then accelerated and mass separated in a magnetic field. Figure adapted from www.cern.ch/ISOLDE/ and Deicher [2]

2.2 In-Flight Techniques

In the in-flight method an energetic heavy ion beam impinges on a thin target and the resulting fragment beam is immediately transported to an in-flight isotope separator. Heavy ion beams at several tens of MeV/u or higher are used [4]. A cocktail of many different species is produced which, since the ions have high velocities, does not need further acceleration to transport it to the secondary target. The activities, projectile fragments, are produced almost at rest in the projectile frame, i.e., produced as a beam well focused in angles and velocities in the laboratory frame of reference. En route to the target the reaction products can be identified by mass, charge, and momentum in a spectrometer (fragment separator). Thus a pure beam is not separated out from the cocktail; instead each ion is tagged and identified by these primary characteristics and the secondary reactions are studied on an event-by-event basis. The in-flight method used for radioactive ion beam production is shown schematically in Fig. 2.

In-flight facilities can produce all chemical species with half-lives greater than about 150 ns, i.e., the time of transit through the fragment separator. The main drawbacks of this method are that the beams are weak, they are not separated physically (the individual ions are simply tagged electronically by A, Z, and momentum) and the beams are of poor quality in terms of energy and focussing. An advantage of
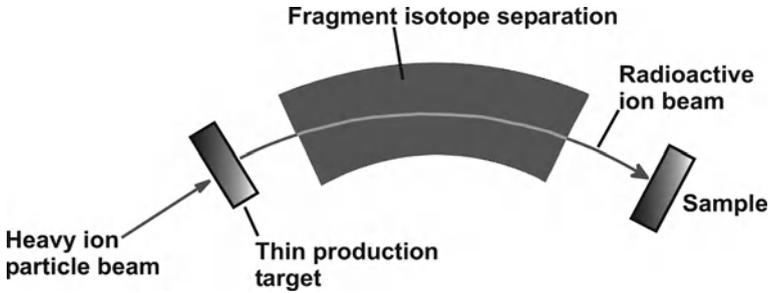
Fig. 2 Schematic view of the basic in-flight method

this rather new technique compared with the ISOL method is that it is chemistry-independent in nature.

A recent significant improvement at the European particle physics laboratory, Cern/ISOLDE in Geneva, is the REX-ISOLDE facility enabling radioactive isotope implantation also at high energies. The exploitation of this facility in materials research is emerging. More than 600 different isotopes of 64 elements can be produced as on-line beams. Given the ferment of ideas and activity in this field it is not surprising that there are many proposals to build new facilities and to upgrade the existing major facilities, e.g., those at GSI (Germany), GANIL (France), RIKEN (Japan), TRIUMF (Canada) and MSU (USA). Lists of existing radioactive nuclear beam facilities, as well as of those planned or under construction, may be found in [5].

3 Materials Characterization with Radioactive Ion Beams

In the following, the most common techniques based on radioactive ion beams used for probing materials at the nano- and even atomic scales are described. Such beams provide analytical tools capable of gaining microscopic information on the structural and dynamic properties of materials. After doping of the material with radioactive probe atoms via ion implantation, as the most versatile procedure, the emission of radiation following decay or the hyperfine interactions, i.e., interaction of nuclear moments with fields in solids, can be exploited. In this way, information about the diffusion processes, lattice sites, and interactions with defects of the probe atoms are obtained, as well as about the internal electric and magnetic fields in crystals and thin films. The procedure of deducing local information on the nature and surroundings of the sites the radioactive probe atoms occupy is illustrated in Fig. 3. It should be emphasised that the technique is essentially a bulk technique that provides structural information on a subnanometre scale and, because of its high sensitivity, can be used to probe interfaces and nanometre films.

Although there has been some emphasis to date on studies of semiconductors, such techniques have also been used to study metals, insulators, superconductors,
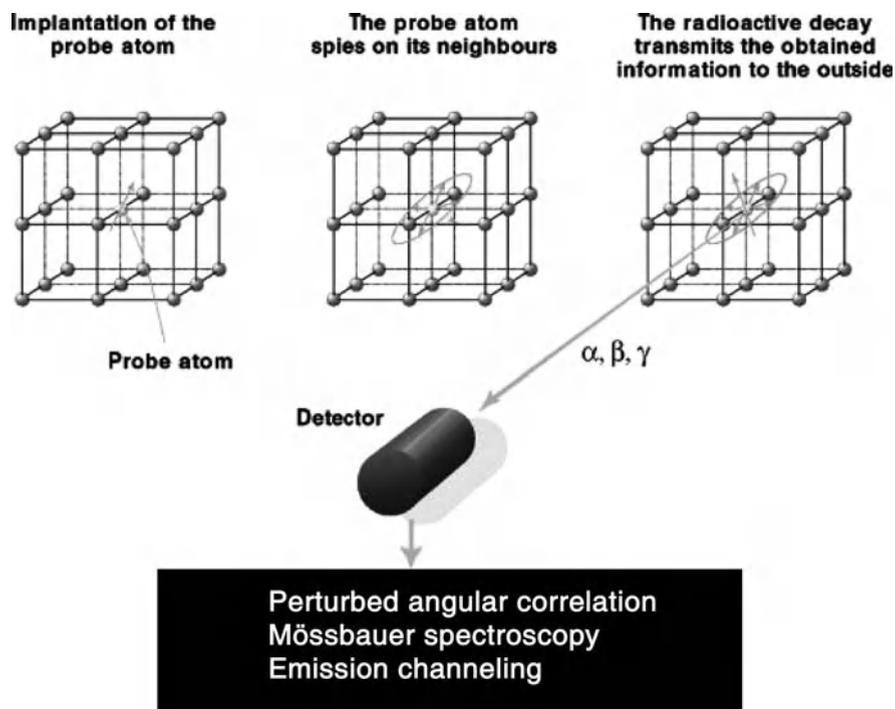
Fig. 3 The properties of the implanted probe atoms are modified by the local microscopic environment. The information of interest is obtained by recording the radiation emitted because of radioactive decay. Figure from Gelletly [6]

surfaces, interfaces, and even complex biomolecules. Because these techniques commonly employ ion implantation, the general possibilities/restrictions connected with ion implantation in stable nuclei are valid. For nanoscale material modification, for example, extremely low ion energies can be used, producing shallow depth profiles in the studied materials. In this way information may be obtained even from the top few nanometre-thick layers. By varying the ion energy, the radioactive probe atoms can be placed at some desired depth. For even more versatile applications of radioactive probe atoms, various lithography-etching techniques commonly employed in research aiming at, e.g., nanotechnological devices, could be combined and taken advantage of. But it should be stressed that this is still an almost fully unexploited region of research, and no detailed studies aiming in these directions have been presented in the literature.

3.1 Radiotracer Diffusion and Emission Channeling

In the following, the so-called modified radiotracer technique for studying solid state diffusion is described and the potential of its use for nanoscience needs is sur-

veyed. In the second part, the emission channeling technique employed for impurity atom lattice site determination is described, and recent prominent experiments are provided as typical examples of its use.

3.1.1 Radiotracer Diffusion

In the development of new materials for a wide range of technological applications, understanding material diffusion properties has played a key role. The conventional method of performing radiotracer diffusion studies has been employed since the 1920s [7]. The procedure of the modified radiotracer incorporates nuances enabled by the development of modern experimental techniques: radioactive ion beams followed by ion implantation and ion beam sputtering.
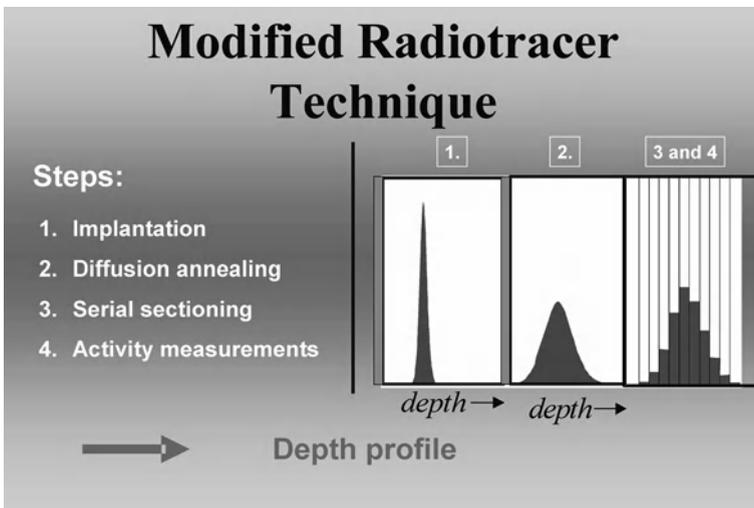
Fig. 4 Experimental steps in the modified radiotracer technique

The basic principles of the modified radiotracer technique are presented in Fig. 4. The technique includes the following steps: (i) production and implantation of the radioactive isotopes of the diffusing element, the tracers; (ii) sample annealing; (iii) serial sectioning by ion beam sputtering followed by measurements of the activity of the eroded material and the depth profile construction; and (iv) quantitative analysis of the experimental depth profile.

A tracer with a convenient half-life and decay mode is first selected. In practice, the isotopes with the shortest half-life employed in diffusion experiments so far are ^{8}Li ($T_{1/2} = 0.84$ s) [8] and ^{11}C ($T_{1/2} = 20.3$ min) [9]. Implanting the tracers just below the material surface allows the influence of surface layers, e.g., oxides, to be minimized. With improved depth resolution, the use of ion sputtering provides possibilities to carry out diffusion studies at lower temperatures than with conventional techniques. In the activity measurements special attention should be paid to efficient

radiation detection, as this decreases the required amount of implanted ions and/or reduces the measurement time.

As an example of the potential of the modified radiotracer technique for studying diffusion, the related results for self-diffusion in $Si_{1-x}Ge_x$ alloys, which are important electronic materials (see chapter "High Speed Electronics"), are presented. The method allows experiments under intrinsic conditions because of its superior sensitivity; furthermore, self-diffusion studies are not feasible with most experimental techniques. Examples of depth profiles obtainable by the technique are shown in Fig. 5, where the measured profiles for ^{31}Si in $Si_{0.20}Ge_{0.80}$ alloy before and after diffusion annealing are presented [10]. Note the extremely low concentrations detectable by the technique.

Fig. 5 ^{31}Si profiles in $Si_{0.20}Ge_{0.80}$ epilayer produced by implantation and subsequent diffusion annealing. The diffusion coefficient at the specified temperature is obtained by fitting the appropriate solution of the diffusion equation to such profiles, which is shown by the *solid line*
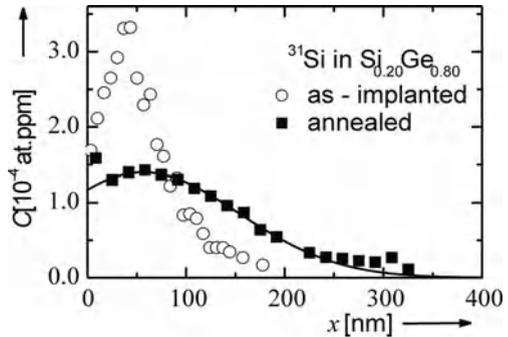
The diffusion properties of self- and impurity atom diffusion in Si and Ge are different. In this light, it is expected that the diffusion properties of $Si_{1-x}Ge_x$ alloys will experience changes as a function of composition. The questions of which compositions, and how these changes happen, have been of major interest.

Potential of the Radiotracer Technique in Nanoscience

The use of the modified radiotracer technique for studying diffusion is currently not very widespread and only a few laboratories have the potential for such experiments. In the following, the feasibility of the technique for profiling depths at the nanometre scale is outlined. Studying diffusion on the nanoscale is a challenging task, as usually experiments are carried out on thin film or multilayer samples of a few nanometre thickness employing techniques such as Auger depth profiling [11]. Using ion beam sputtering, sectioning of even monolayers is feasible, which provides possibilities for clearly improved depth resolution. The sputtering rate of the studied material affects the collection times. To ensure sufficient accuracy in the sputtering time measurement, the ion energy should be selected at a properly lower level. The other factor that should be taken into account is the sensitivity of the activity measurement. For alpha particles, electrons, and positrons, the detection efficiency can be easily maximized. The isotope half-life also plays a role. Typically, depth profiling at a resolution of about 0.4 nm is achievable for an isotope with

1–4 h half–life, when implantation fluences on the order of 10^9 cm^{-2} are employed. Factors playing a role and hindering a better depth resolution are, e.g., purity of the surface (possible oxide/impurity layer), homogeneity of the beam used for sputtering, and physical processes taking place at the sputtered surface (e.g., preferential sputtering). It should be noted that employing shallow implantation and using etching techniques and diffusion on the lateral nm-scale could in principle be studied.

3.1.2 Emission Channeling (EC)

The sensitivity of the conventional RBS/channeling technique (see chapter "Thin Film Characterisation by Means of MeV Ion Beams") can be significantly improved by using as probes implanted radioactive impurity atoms that emit charged particles, so-called emission channeling. The technique dates back to the mid 1960s when it was introduced [12], and it provides information even to an accuracy of a few hundredths of a nanometre on the lattice sites of radioactive isotopes in single crystalline materials. The process is described schematically in Fig. 6.
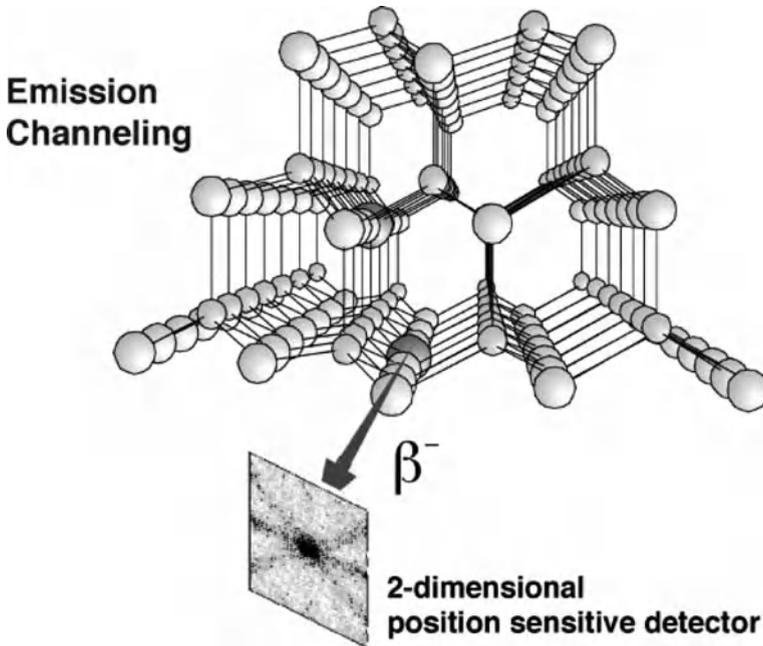
Fig. 6 Emission channeling of electrons detected with a 2-dimensional position sensitive detector. Figure from Deicher [2]

The yield of the charged particles (β^-, β^+, conversion electrons, α- particles) emitted by the implanted radioactive probes is recorded for different emission directions with respect to the crystallographic axes or planes of single crystals. Depending on the lattice site of the emitting nuclei and the nature of the emitted particles,

they experience either channeling or blocking effects. For example, if the probe atom emitting positively charged particles is on a substitutional lattice site, the positively charged particles will be blocked by the atom rows and planes, leading to large-angle scattering (de-channeling). This results in minima in the emission yields along the axes or planes (blocking). If the atom is located on an interstitial lattice site, more favorable emission directions exist along which the particles will be guided in a series of small-angle scattering processes. Along such directions enhanced emission (flux peaking) will be noted. In the case of β^- and conversion electrons, an enhanced emission yield along a lattice direction (channeling) results in an emitting atom residing on a substitutional site. A reduced yield or a weak anisotropy (blocking) points to an interstitial lattice site of the radioactive probe. The theoretical treatment of emission channeling (EC) in the case of electrons is more complicated, requiring both relativistic and quantum mechanical calculations. In the case of alpha-particles and positrons, the procedure is based on classical models using Monte Carlo simulations.

The detection efficiency for the angular distribution of the emitted particles is increased by using two-dimensional position sensitive detectors, as shown in Fig. 6. This enables the use of very low implanted fluences, on the order of $10^{12}\,\mathrm{cm}^{-2}$ [13], which is about two orders of magnitude lower than the fluences needed in conventional RBS/channeling. The EC technique has been applied in dopant atom lattice location studies of elemental semiconductors, numerous compound semiconductors, metals, diamond, and several oxides. For recent emission channeling studies on lattice location of Fe and rare earth elements in the wide band-gap semiconductors GaN and ZnO, the reader should consult [14].

Potential of the Emission Channeling Technique in Nanoscience

It should be pointed out that a very limited number of research groups carry out EC experiments, and all of them are interconnected to the Cern/ISOLDE facility. As ion implantation of the probe atoms is the key feature in EC experiments, the valid possibilities for ion implantation in general hold true also for EC-based research. Shallow implantation and material etching offer possible additional tools for carrying out nanoscience related studies with EC.

3.2 Defect Configuration and Lattice Location of Impurities via Hyperfine Interactions

The radiation pattern of implanted probe nuclei is influenced and altered by the electric and magnetic fields induced by the surroundings of the probes in the host material. By monitoring these changes, information on the nature of their local atomistic environment can be obtained. The introduced techniques of Mössbauer spectroscopy, perturbed angular correlation, β-NMR and nuclear orientation belong to this category of nuclear hyperfine techniques. They are based on the observation of the electric and magnetic hyperfine interaction between the nuclear moments of

specific probe nuclei and the magnetic or electric fields induced by the surroundings of the probes (valence electrons, conduction electrons, neighbouring atoms or ions, neighbouring defects). The hyperfine interaction is very short-ranged and is determined by both the nuclear wave functions and the tails of the electron wave functions that reach the nucleus. Therefore bonding effects on the valence electrons caused by the atom being contained in a molecule, liquid, or solid are transferred to the nuclear hyperfine structure. Hence, measurements of hyperfine structure provide information on the chemical environment of the atom. The high sensitivity combined with the small number of probe atoms required (up to 10^{13} atoms in total) make them ideal tools for studying structural, electronic, and dynamical properties of self-atoms and single impurities, and of inter-impurity and impurity-defect complexes on an atomic scale. These techniques provide material probing at the smallest possible length scale. As the above-mentioned emission channeling technique provides information on the lattice sites of the impurity atoms, these hyperfine methods provide information on the local environment of the probe atoms, and they all can be applied in a complementary way. Several comprehensive review articles dealing with hyperfine interaction basics, theoretical treatment, and applications can be found in the literature.

3.2.1 Mössbauer Spectroscopy (MS)

The basic idea of the well-known Mössbauer effect is the recoil-free emission of gamma quanta from radioactive nuclei embedded in a solid. If the emitted gamma quantum populates the ground state of the nucleus, it is possible to resonantly absorb a recoil-free emitted gamma quantum (source) by the same nucleus in its ground state (absorber). To get a reasonable sensitivity, the system under study acts as the host matrix of the source atoms and is compared to a standard absorber. The electric charge distribution around the probe nucleus creates a shift of the resonance energy due to the electric monopole interaction, giving rise to the isomer shift, which is only accessible by the Mössbauer effect. The information that can be extracted from the isomer shift is the lattice site of the Mössbauer atom and its charge state. The latter parameter indicates, for example, whether the probe atom in a semiconductor acts as a donor and has emitted its electron into the conduction band. The technique provides thus detailed chemical and structural information on the environment at the Mössbauer nucleus. It is generally considered that Mössbauer spectroscopy is one of the most sensitive tools for studying the structural and magnetic properties of materials.

Figure 7 illustrates how the material structure alters the Mössbauer spectra [15]. Crystalline materials feature discrete, narrow, and well-separated spectral lines. They reflect precise resonant atom positions in the well-defined crystalline material, and distinct crystalline phases can be unambiguously identified from such spectra. If the spectrum is doublet-like, the studied material is non-magnetic, whereas a sextuplet feature is characteristic of magnetic materials. Amorphous materials exhibit broad and overlapping spectral lines, but also in this case a distinction between magnetic and nonmagnetic materials is possible.
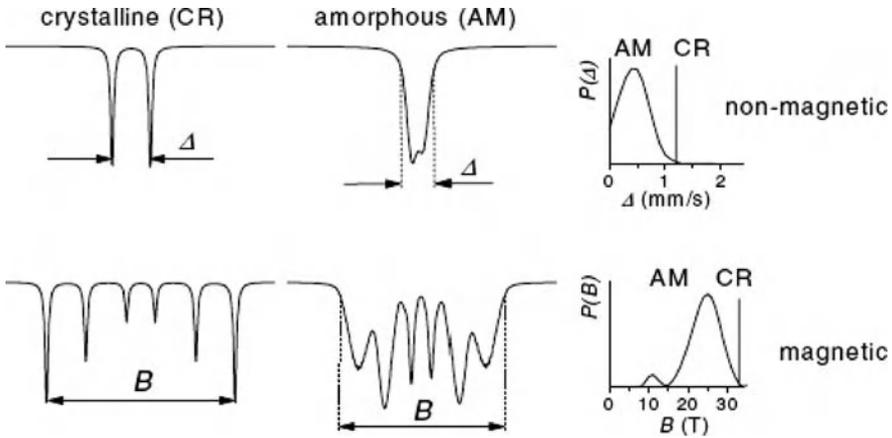
Fig. 7 Mössbauer spectra of crystalline and amorphous materials in nonmagnetic and magnetic states. Figure has been adapted from Miglierini [15]

Until very recently Mössbauer spectroscopy has been of limited use. This is because it relies on the existence of pairs of nuclei such as ^{57}Fe/^{57}Co, where a low energy Mössbauer state of a stable isotope is fed in the decay of a long-lived parent nucleus which can be conveniently produced, purified, and incorporated in a suitable matrix. If the parent nuclei are produced as a radioactive beam and using ion implantation, the requirement for a long half-life can be left aside. Under such circumstances a large number of new Mössbauer systems can be identified, totalling about 75 cases among over 40 elements. However, at the current time, over 65% of the Mössbauer studies are carried out employing the ^{57}Fe isotope.

Mössbauer spectroscopy has been recently employed to study the relationship of nanocrystalline alloy magnetic properties to their microstructure [15]. The macroscopic parameters of nanocrystalline alloys are superior to conventional and amorphous alloys because their magnetic properties can be tailored according to needs. In such studies the merits of Mössbauer spectroscopy are invaluable because the structural characteristics and magnetic interactions can be probed simultaneously. An example of the use of MS for studying diffusion at the true microscopic scale is the study of [16]. In this work radioactive ^{57}Mn isotopes were implanted into $Si_{1-x}Ge_x$ alloys at elevated temperatures, and the diffusion jumps of interstitial ^{57}Fe daughter atoms were followed. In the study it was shown that the atomic jump frequency increases as a function of Ge alloying.

3.2.2 Perturbed Angular Correlation (PAC)

The PAC technique is very closely related to Mössbauer spectroscopy. The magnetic and quadrupole hyperfine interaction is detected by observing the energy splitting of the m- substates of the isomeric state of a suitable probe atom. PAC measurements can be performed at room temperatures up to the melting point of the host lattice and even beyond. The experiments require probe atoms which decay by a
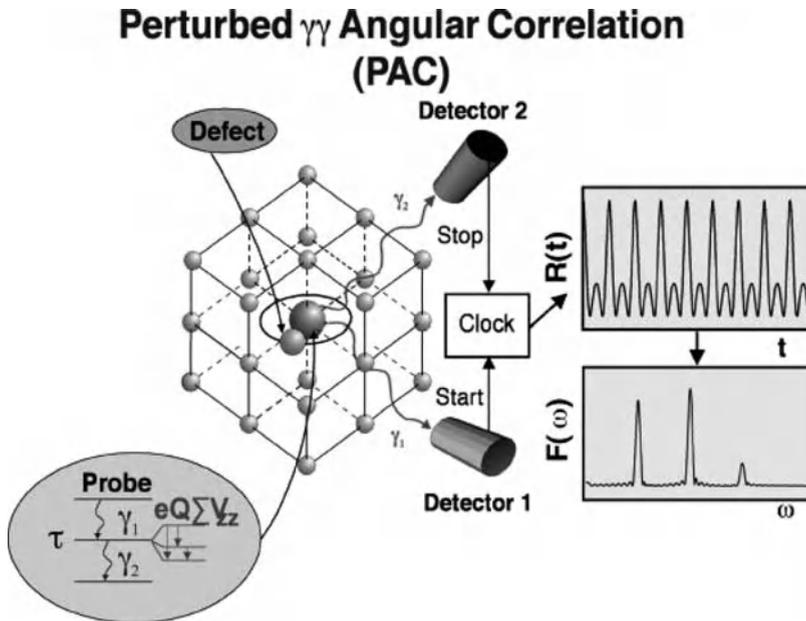
Fig. 8 Principles of the PAC method. If the PAC probe atoms are localized on highly symmetric lattice sites in a cubic or diamond structure, the electric field gradient is zero and the coincidence ratio $R(t)$ shows no modulation as a function of time. However, as shown in the figure, when a defect or impurity is present and disrupts the symmetry around the probe atom, a time-modulated $R(t)$ spectrum is obtained (shown to the *right*). Figure has been adapted from Deicher [2]

$\gamma\gamma$- cascade via an isomeric nuclear state. By measuring the two γ- quanta (γ_1, γ_2) in coincidence, an anisotropic angular distribution of the γ_2 radiation is observed (see Fig. 8). If the charge distribution around a nucleus in a lattice has a symmetry lower than cubic, an electric field gradient is generated at the site of the nucleus. This situation exists in noncubic crystal lattices and when defects are present in the neighbourhood of the probe atom. The electric field gradient provides information about the symmetry and orientation of the charge distribution with respect to the crystal axis. This electric field gradient is easily measurable using the PAC technique. The PAC technique is more versatile compared to other hyperfine techniques because the choice of suitable isotopes is wider and the half-lives of the parent probe isotopes can be on the order of minutes to days. Using the PAC technique it is possible to discriminate the presence of extrinsic (impurity atoms) from intrinsic (vacancies/self-interstitials) defects. This is of fundamental interest, e.g., in defect related physics.

About 10^{11} probe atoms are needed for reliable PAC experiments. The probe atoms are commercially available, produced at a reactor or by a particle accelerator. The most widely used probe nuclei are ^{111}In, ^{181}Hf, ^{100}Pd, ^{77}Br, ^{117}Cd, ^{99}Ru, ^{197}Hg, and ^{44}Ti. The PAC sample can then be prepared by conventional alloying procedures or by implantation of the radioactive probe atoms into the material.

As an example of the potential of the PAC technique, the case of hydrogen in semiconductors can be demonstrated. Hydrogen represents one of the most important impurities in semiconductors, both from the scientific and technical points of view. It is easily incorporated during manufacture and it interacts very efficiently with other impurities or defects. The free migration of well-defined H configurations in III–V semiconductors such as InP can be studied on an atomic scale using PAC with the probe nucleus ^{117}Cd, which is populated in ^{117}Ag decay. The process is started by irradiating the sample with a pure ^{117}Ag beam. After annealing, the sample is loaded with H$^+$. The Coulomb interaction between ionised acceptors (Cd$^-$) and donors (H$^+$) induces the formation of Cd-H pairs. After the decay of ^{117}Cd into ^{117}In the pair is transferred to In-H, and because the H$^+$ donors no longer feel the attractive Coulomb force they are able to migrate. The onset of this migration process can be observed by PAC measurements [17].

The field of ultrathin metallic layer magnetism has become of significant interest. The PAC technique ideally provides information on the magnetic property variation at the atomic layer level or even at the monolayer level. Recent PAC studies of ferromagnetic semiconductors using different nuclear probes produced at ISOLDE have been summarized in [18]. In [19] a summary of recent PAC studies related to magnetic properties of impurities at ferromagnetic surfaces and interfaces may be found. It may be noted that the number of PAC studies related to nanocrystalline semiconductors is clearly expanding. An example of such a PAC study is the determination of the structural properties of nanocrystalline ZnO doped with ^{111}In [20].

3.2.3 β-NMR

Nuclear Magnetic Resonance (NMR) and related methods are widely used in condensed matter studies. The magnetic moment of the probe nucleus acts as a sensitive probe of the local magnetic and electronic properties of materials. Generation of a non-equilibrium nuclear spin polarization followed by the temporal behaviour of the polarization is required. In traditional NMR about 10^{18} nuclear spins are needed to generate good NMR signals, and therefore NMR is used mostly in studies of bulk materials.

The β-*NMR* technique is the radioactive version of the NMR, where the signal comes from the decay properties of a radioactive nucleus. This procedure possesses an improved sensitivity of a factor of several orders of magnitude. Since now a factor of 10^{11} fewer spins are needed compared to the conventional NMR case, this technique is well suited for studies of dilute impurities, nanostructures, and thin films. Polarized β- emitters are either directly produced in the lattice via nuclear reactions with polarized thermal neutron beams, or they are implanted after production and polarization. To preserve the polarization of the probe atoms, the host lattice has to be located in a magnetic field which is oriented parallel to the direction of polarization. The hyperfine interaction is detected via the β-asymmetry of the decay radiation of the implanted radioactive probe atoms.

Although in principle any β-emitting isotope with a spin can be employed in β-NMR, the number of isotopes suitable for condensed matter studies is clearly

constricted. The requirements for the probes are: (1) high production efficiency, (2) efficient polarization of the nuclear spins, and (3) high β decay asymmetry. It is also advantageous if the atomic number of the probe atom is low, in order to minimize radiation damage during implantation. The probe should possess a small value of spin to assure simple β-NMR spectra. Furthermore, the half-life of the probe should not be much longer than a few seconds. The ^{8}Li isotope is the easiest probe to polarize, but other suitable isotopes are, for example, ^{11}Be, ^{12}B, ^{12}N, ^{15}O, and ^{17}Ne.

At the β-NMR facility of ISAC, TRIUMF, low energy, highly spin-polarized ^{8}Li ion beams are used for studying thin films and nanostructures [21]. As the nuclear spins of the implanted probe nuclei are affected by the surrounding distribution of nuclei and electrons in the sample material, a good sensitivity to the surrounding electronic and magnetic properties is achieved. An advantage of this method is the possibility of controlling the implantation depth within the scale (1–400 nm). This technique has been employed in studies of multilayers of magnetic/nonmagnetic conducting layers similar to the read/write heads of computer hard drives; see Fig. 9. With low implantation energy most of the Li- ions stop in the top 4 nm thick gold layer, and the resulting resonance signal is shown on the right. By increasing the implantation energy, the ^{8}Li probes are stopped in the
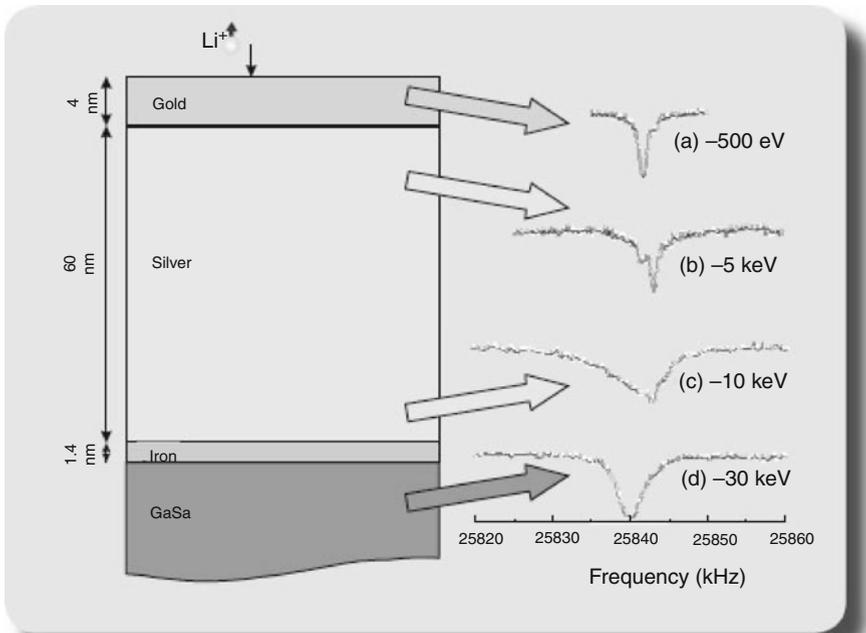
Fig. 9 An example of the use of β-NMR where ^{8}Li ions are implanted to various depths of a multilayered structure deposited on a GaAs substrate. The figure is from TRIUMF Annual Financial and Administrative Report 2005 [22]

desired layer, allowing measurements of their internal magnetic fields. In this way the technique becomes unique and powerful for depth-resolving measurements in ultra-thin films and nanostructures.

3.2.4 Nuclear Orientation (NO)

In this method, the polarization or alignment of the nuclear spins of the probe atoms must also be achieved to create an anisotropic emission of the decay products. In this technique, the probe atoms are implanted into a ferromagnetic crystal (e.g., Co, Fe, Ni), which is then cooled down to temperatures of few mK. This technique is mostly employed in nuclear physics for studying nuclear moments and excited states of rare isotopes.

Potential of the Hyperfine Techniques in Nanoscience

A better understanding of the interactions between different magnetic multilayers is essential to improve the sensitivity of, for example, magnetic memory heads. Specifically, such information is achievable by employing various hyperfine interaction techniques.

At the moment there are roughly a dozen research groups in the world applying hyperfine interaction techniques for studying local structures and electronic properties at surfaces and in thin film multilayers. The short range of the hyperfine interactions allows monolayer resolution for layered systems and in some cases even an atomic resolution. The long penetration range of the radiation detected enables measurements of buried layers and, combined with the ability to adjust the beam energy, offers unique advantages for "focusing-in" on nanostructures. This allows measurements of structures as thin as 1 nm, or alternatively, the study of different sample layers by adjusting the beam energy so that the ions stop at a specific desired depth. To date, the full potential of the hyperfine methods has not been utilized in nanoscience, but the activity is clearly moving in this direction, as illustrated by the International Workshop on "Nanomaterials, Magnetic Ions, and Magnetic Semiconductors Studied by Hyperfine Interactions" held in 2004 [23]. Because with these techniques the local properties on a truly atomic scale can be determined, the aims of miniaturisation of electronic devices may be set in an entirely new light.

4 Standard Electrical and Optical Techniques Combined with Radioactive Isotopes

Many conventional methods, such as photoluminescence (PL), capacitance voltage (CV), the Hall Effect, and deep level transient spectroscopy (DLTS), etc., are not capable of identifying the chemical nature of the defects they are detecting. Combining these techniques with the radioactive isotope techniques, high sensitivity for low defect concentrations and chemical nature information is achieved. The

intensity of the observed electrical or optical signals changes with a characteristic time constant based on the radioactive decay. This allows one to assign the electrical or optical properties to a particular chemical element unambiguously. For example, with a known radioactive probe half-life and observation of the defect concentration change, and following its time dependence, the lacking chemical information for the data obtained by PL can be obtained. Many such recent experiments have dealt with the new semiconductor materials of interest in optoelectronics, especially GaN [24]. The trend to combine such conventional techniques with modern radioactive probing techniques is at present significantly increasing.

References

1. Piercy GR, Brown F, Davies JA, McCargo M (1963) Experimental evidence for the increase of heavy ion ranges by channeling in crystalline structure. Phys Rev Lett 10: 399–400.
2. Deicher M (2002) Radioactive isotopes in solid state physics. Europhys News 33.
3. Dendooven P (1997) The development and status of the IGISOL technique. Nucl Instrum Methods B 126: 182–189.
4. Asahi K, Kobayashi Y, Ueno H, Yoshimi A, Sato W, Watanabe H, Miyoshi H, Kameda D (2001) Radioactive ion beams as microprobes into matter. Anal Sci 17: 613–616.
5. http://www.ganil.fr/eurisol/EURISOLlinks.html
6. Gelletly W (2001) Science with radioactive beams: the alchemist's dream. Contemp Phys 42: 285–314.
7. Hevesy GV (1920) Die Platzwechselgeschwindigkeit der Ionen im Kristall. Z Phys 2: 148–149.
8. Jeong SC, et al. (2005) Measurement of diffusion coefficients in solids by the short-lived radioactive beam of ^{8}Li. Nucl Instrum Methods B230: 596–600.
9. Voss T, Strohm A, Frank W (2003) Diffusion in polymer-derived Si-(B-)C-N ceramics, particularly amorphous $Si_{29}B_9C_{41}N_{21}$. Z Metallkd 94: 419–423.
10. Laitinen P, Strohm A, Huikari J, Nieminen A, Voss T, Grodon C, Riihimäki I, Kummer M, Äystö J, Dendooven P, Räisänen J, Frank W, the ISOLDE Collaboration (2002) Self-diffusion of ^{31}Si and ^{71}Ge in relaxed $Si_{0.20}Ge_{0.80}$ layers. Phys Rev Lett 89: 085902-1–085902-4.
11. Beke DL, Erdélyi Z, Langer GA, Csik A, Katona GL (2005) Diffusion on the nanometer scale. Vacuum 80: 87–91.
12. Domeij B, Björkqvist K (1965) Anisotropic emission of α-particles from a monocrystalline source. Phys Lett 14: 127–128; Uggerhøj E (1966) Orientation dependence of the emission of positrons and electrons from ^{64}Cu embedded in single crystals. Phys Lett 22: 382–383; Astner G, Bergström I, Domeij B, Erikson L, Persson A (1965) Angular dependence of conversion electrons from a monocrystalline source. Phys Lett 14: 308–310.
13. Correia JG (1996) Radioactive ion beams for solid state research. Nucl Instrum Methods B 113: 501–506.
14. Wahl U, Correia JG, Rita E, Alves E, Soares JC, DeVries B, Matias V, Vantomme A, the ISOLDE Collaboration (2004) Re*-cent emission channeling studies in wide band gap semiconductors. Hyperfine Int 159: 363–372.
15. Miglierini M (2005) How can Mössbauer spectrometry contribute to the characterization of nanocrystalline alloys? Hyperfine Int 164: 41–49.
16. Weyer G, Gunnlaugsson HP, Bharuth-Ram K, Dietrich M, Mantovan R, Naicker V, Naidoo D, Sielemann R (2004) Acceleration of diffusional jumps of interstitial Fe with increasing Ge concentration in $Si_{1-x}Ge_x$ alloys observed by Mössbauer spectroscopy. Hyperfine Int 158: 417–421.

17. Forkel-Wirth D (1999) Exploring solid state physics properties with radioactive isotopes. Rep Prog Phys 62: 527–597.
18. Samokhvalov V, Dietrich M, Schneider F, Unterricker S, the ISOLDE Collaboration (2005) The ferromagnetic semiconductor $HgCr_2Se_4$ as investigated with different nuclear probes by the PAC method. Hyperfine Int 160: 17–26.
19. Manzhur Y, Prandolini MJ, Potzger K, Weber A, Zeitz WD, Bertschat HH, Dietrich M (2005) Surface and interface magnetism using radioactive probes. Hyperfine Int 160: 3–15.
20. Agne T, Deicher M, Koteski V, Mahnke HE, Wolf H, Wichert T (2004) Structural properties of the donor indium in nanocrystalline ZnO. Hyperfine Int 159: 55–61.
21. Kiefl RF, et al. (2003) Low-energy spin-polarized radioactive beams as a nano-scale probe of matter. Physica B326: 189–195; Low energy spin polarized radioactive beams as a probe of thin films and interfaces. Nucl Instrum Methods B204: 682–688.
22. Salman Z (2005) Thin films and nanostructures studied by β-NMR. TRIUMF Annual Financial and Administrative Report 2005: 30–33.
23. Somayajulu DRS, Lieb K-P (eds) (2005) Proceedings of the international workshop on nanomaterials, magnetic ions and magnetic semiconductors studied by hyperfine interactions. Hyperfine Int 160: 1.
24. Dalmer M, Vetter U, Restle M, Stötzler A, Hofsäss H, Ronning C, Moodley MK, Bharuth-Ram K, the ISOLDE Collaboration (1999) Combination of emission channeling, photoluminescence and Mössbauer spectroscopy to indentify rare earth defect complexes in semiconductors. Hyperfine Int 120–121: 347–352.

Part IV
Nanoscale Materials Processing with Ion Beams

The fact that ion beams are electrically charged not only makes them well suited as probes for characterisation, (discussed in the previous part) but also makes possible their use for modification on an nanometre scale. This strong interaction makes possible extremely localised deposition of energy into the electronic and atomic structures of the target – often so the energy is delivered to a single atom.

The extreme localisation of the deposited energy is made use of for proton beam writing lithography and in focused ion beam technology (FIB). These two techniques, although they both use focused ion beams, because of the different energies and mass of the ion beam used, work on quite different physical principles with very different and in many ways complementary characteristics. FIB is a very standard workhorse tool in nanoscience, where it is used as a nanometre-scale material deposition and cutting torch, which makes it useful for modifying prototype integrated circuits by depositing metal and insulating layers as well as cutting out specimens for transmission electron microscopy. Proton beam writing, on the other hand, is capable of rapidly writing nanometre-scale patterns in thick polymer layers and even writing patterns of defects that can be used for tasks as diverse as tissue engineering, prototyping lab-on-a-chip devices, and making colour-tunable optoelectronic devices in nanoporous silicon.

Broad-beam ion bombardment is also very important in nanoscience and technology. Ion irradiation by a plasma forms the basis of plasma etching and deposition. This is a workhorse technique that is used for a range of purposes in nanofabrication, from making 3D patterns in hard materials like silicon by anisotropic etching in combination with masking, to depositing refractory metals like tungsten and platinum, as well as cleaning wafers by "ashing" resist polymers in an oxygen plasma. Most significantly, ion implantation can be used in a bottom-up way to form nanoparticles in materials in which, because of quantum confinement and coupling, they modify the fluorescence and other optical properties of the materials.

Nanocluster and Nanovoid Formation by Ion Implantation

Kai Nordlund and Juhani Keinonen

1 Introduction

Ion implantation is a well-established technique to modify the near-surface region of materials by the introduction of keV–MeV ions. We discuss the nucleation, growth, and modification of nanoclusters and nanovoids as a part of ion implantation techniques in the field of nanomaterials.

A proper choice of the experimental parameters of the implantation process (energy, current, fluence) allows for predetermining the composition, depth, and spatial shape of the modified layer. A fundamental feature of ion implantation is that the implantation process is not a thermodynamic equilibrium process. Consequently, the usual solubility limits of the implanted atoms in the host can be largely overcome, achieving local atom concentrations inaccessible by conventional synthesis routes. The irradiation-induced change of diffusion rates into the solid can favour either the aggregation of the dopant atoms or their diffusion inside the target. Depending on the choice of the implanted atoms and target material, implantation gives rise to the formation of new compounds, nanoparticles, nanoparticle composites, and nanovoids.

The ion implantation and irradiation-induced modifications of materials surface layers are correlated to the interactions of the energetic ions with electrons (electronic energy loss) and atoms (nuclear energy loss) of a target material. The interaction processes occur on a time scale of picoseconds and length scales as low as nanometers. The elastic collisions in the nuclear energy loss produce either highly damaged lattices following implantation or a network of damaged structures whose density depends on lattice reordering during the implantation process, i.e., defect production and dynamic annealing. In insulating materials, radiation damage is induced also in the electronic stopping power regime, due to the breaking of interatomic bonds.

K. Nordlund (✉)
Department of Physics, University of Helsinki, Helsinki 00014, Finland
e-mail: kai.nordlund@helsinki.fi

To understand how nanoclusters and nanovoids can be formed in materials by ion implantation, one must understand the concepts of ion range and deposited energy distributions. Although these have been discussed in preceding chapters of this book (Part II, Box 3), we briefly summarize here these concepts.

The depth distribution of implanted ions is usually called the ion range profile. It typically has a well-defined maximum and a Gaussian shape around this maximum, although the tail and near-surface regions may significantly differ from this shape. The mean of this distribution is called the mean range, and typically it coincides with the maximum of the range distribution.

The distribution of nuclear deposited energy does not exactly coincide with the range distribution. This is because the ions transfer considerable amounts of energy to lattice atoms before they stop. Hence the distribution of nuclear deposited energy has a maximum closer to the surface than the range distribution. Except for this, the two distributions are not very different in shape.

Ion implantation produces atom vacancies and interstitial atoms, either isolated or clustered ones (cf. Part II, Chapter "Computer Simulate Methods for Detect Configurations and Nanoscale Structures and Part III, Box 6). In some materials the implantation may also produce strongly disordered "amorphous" zones, but in the current discussion we assume these recrystallize and leave behind only point defects. This is a reasonable assumption in a wide range of materials implanted at room temperature.

The interstitials produced are typically more mobile than vacancies. Some of them recombine with vacancies, but some can escape the damaged region to the surface and in essence vanish there. The surface is often assumed to be a "perfect sink" for interstitials, and that this is in fact the case has recently been proven for crystalline Si [1]. Hence the net effect of the irradiation is that the implanted zone becomes vacancy-rich.

Figure 1 illustrates an ion-implanted region in a sample. The range and deposited energy profiles were obtained in molecular dynamics (MD) calculations and are compared with an experimental range profile.

The slowing down of the implanted ions occurs until they have lost all their initial kinetic energy and are left with only a kinetic energy corresponding to the temperature of the lattice, i.e., the implanted atoms have become thermalized. The time scale of the process is roughly one picosecond.

2 Formation of Nanoclusters

After the atoms have thermalized, they may be mobile because of ordinary thermal mobility, namely diffusion. In fact, such mobility is a prerequisite for formation of nanoclusters embedded in a lattice. If the implanted atoms are not mobile, they will be distributed randomly in the implanted region of the matrix, and cannot form nanoclusters at reasonable concentrations. But mobility by itself does not suffice for nanocluster formation. If the implanted atoms are miscible with the matrix material,
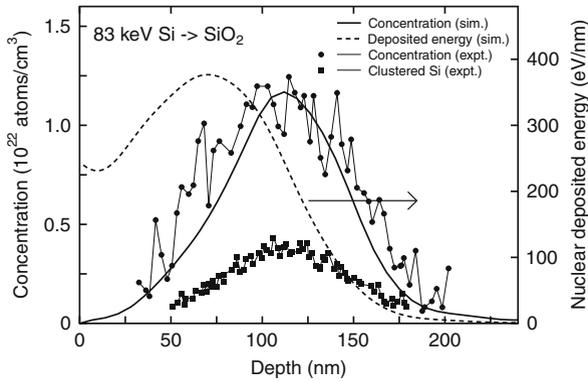
Fig. 1 Illustration of the relation between implanted ion and deposited energy depth distributions. The *full line* shows the ion depth distribution for 83 keV Si ions implanted into SiO_2 simulated with the MDRANGE method [2]. The *dashed line* shows the corresponding deposited energy distribution. Also shown are recent experimental results for the same implantation obtained with transmission electron microscopy [3]. The *circles* show the excess (i.e., implanted) Si depth distribution in the SiO_2, and the *squares* the depth distribution of Si nanoclusters after annealing at $1100°C$ for 1 h. The experimental data was obtained by digitizing the data in Fig. 3b in [3]

they will still tend to be randomly distributed in it, since there is no thermodynamic driving force for segregation. Moreover, atoms miscible with the lattice are not likely to have a significant mobility because they can find energetically favourable locations in the lattice.

If the implanted atoms are immiscible, however, they will naturally tend to segregate from the lattice. Then whenever mobile species of the implanted atom type "find" each other, they will tend to nucleate, i.e., cluster, with each other, and can thus form nanoclusters in the material.

The sample temperature obviously has a major effect on whether and how the clustering occurs. Clearly the temperature must be high enough for implanted atom migration to enable clustering. Typically room temperature is not high enough, and to facilitate the clustering the implanted samples are heated after the implantation. On the other hand, the temperature may also affect the miscibility of the material: some materials may be immiscible at low temperatures, but miscible at higher ones. It is well known that although a binary alloy is completely immiscible at 0 K, it always becomes somewhat miscible at higher temperatures. The *solvus* line in a phase diagram shows where this occurs. Because the *solvus* line typically involves only low concentrations, so does the formation of nanoclusters. Hence a heating used to form nanoclusters should not be so high as to cross the *solvus* line, since in that case no further segregation and cluster formation can be expected to occur.

It is interesting to consider how the cluster formation can be initiated. The simplest possible case is if it is possible for individual impurity atoms A to form a dimer A_2, then a trimer A_3, and so on, with the growth always being energetically favourable [4]. In this case the nanocluster growth can be initiated anywhere in the lattice as soon as atoms A happen to migrate next to each other. However, in many

cases it is not energetically favourable for two atoms A to join together. This may be, e.g., because of a large strain energy for a dimer A_2, or that the chemical bonding energetics is unfavourable. In this case the cluster growth may still be possible as a "nucleation and growth" type of process where a nucleation energy barrier needs to be exceeded before the growth can proceed. However, such a process would obviously be quite slow if the energy barrier is high. On the other hand, most types of ion irradiation produce vacancies in the sample, as discussed above. Cluster formation can start easily in a vacancy, or agglomerate of many vacancies, if several implanted atoms A fit in there.

The structure of the nanocluster formed may be crystalline or amorphous, or even gaseous in the case of noble gas clusters (see e.g., [5–7], and references therein). Predicting the structure is not trivial, however, since the clusters are under a high pressure and have a high surface area. Thus their structure may be considerably different from that in normal equilibrium. A dramatic example of this is Xe nanoclusters, which can be crystalline at room temperature even though the boiling point of Xe at zero pressure is only 165 K [8] (see Fig. 2).

As an example of the processes just discussed, we consider recent studies of H and He bubble growth in W. Using a combination of different atom-level simulation methods, Henriksson et al. [9] found that He bubble nucleation can start even in the perfect W lattice, since it is energetically favourable for two isolated He atoms to form a He pair in the W lattice. H atoms, by contrast, cannot form a pair, and hence initiating bubble nucleation is unlikely. A single vacancy can bind a few H atoms, but not more. Hence bubble nucleation of H atoms in W lattice can only begin at larger defects than a single vacancy. The precise nature of what kind of a defect can initiate the nucleation is hitherto unknown.

Once an initial seed for a cluster has formed, its growth requires that some lattice atoms leave the cluster region. This can be achieved by nucleating an interstitial atom at the interface between a forming bubble and the surrounding lattice. Since interstitials tend to be highly mobile in a wide range of materials, they
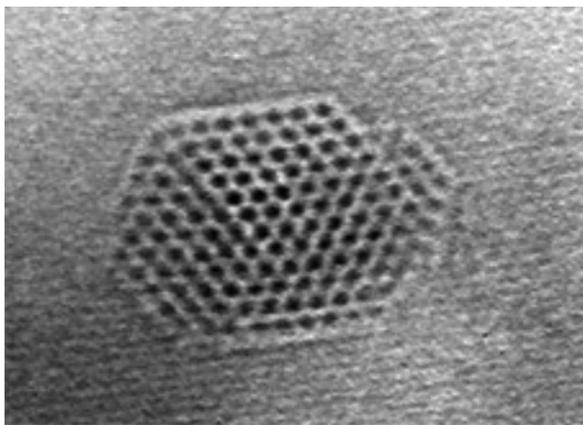
Fig. 2 Solid Xe at room temperature inside an Al matrix. Reprinted with permission from Birtcher et al. [8]. Copyright (1999) by the American Physical Society

can easily leave the cluster region, leaving behind a vacancy which gives "more room" for the cluster to grow. When the growing cluster is large enough, it may be possible to "punch out" an entire interstitial dislocation loop [11]. Although the formation of interstitials normally requires lots of energy, the high pressure associated with a nanocluster or -bubble can reduce the formation enthalpy sufficiently to enable interstitial creation even at moderate temperatures. This mechanism was first treated in detail theoretically by Trinkaus, who predicted for He atoms that 5–7 atoms was enough to cause interstitial nucleation [12]. This prediction has turned out to be remarkably accurate when compared to recent MD simulations [13].

Even if the nanocrystals are crystalline, they can contain a large fraction of defects. For the case of Xe nanoclusters, the motion of dislocations inside the nanoclusters has been directly observed in transmission electron microscopy (TEM) experiments [8].

Once cluster growth has initiated in a temperature range where the cluster-forming ions are mobile in the lattice, it will proceed until all ions are part of a cluster or have escaped to the surface or into the bulk. This process results in a distribution of cluster sizes. The biggest clusters are likely to be located at about the maximum of the ion range distribution, where the deposited energy and thus damage has also a maximum. At the tails of the distribution the cluster sizes are smaller. Hence there is a wide variation in cluster sizes, which may not be desirable for practical applications. Note that the concentration distribution of atoms in the clusters still correlates with the original range distribution, see Fig. 1. Figure 3 illustrates the process of cluster growth.
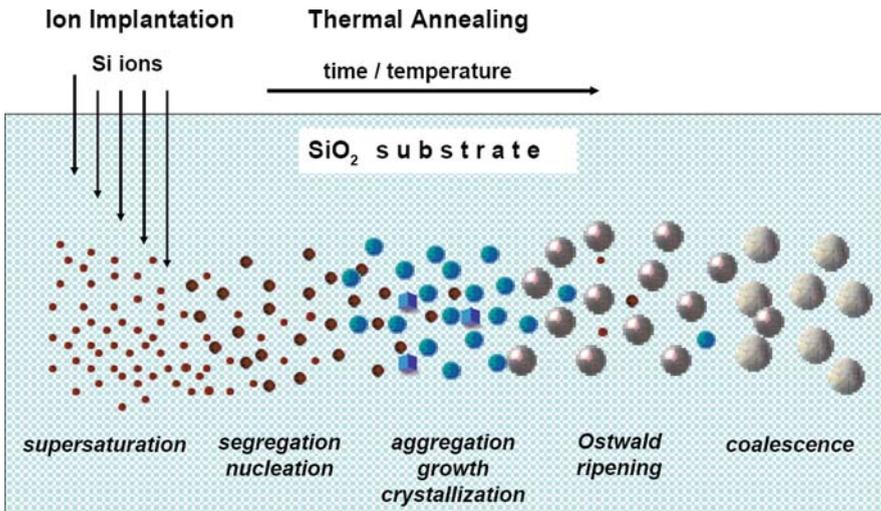
Fig. 3 Schematic illustration of the process of cluster growth and Ostwald ripening
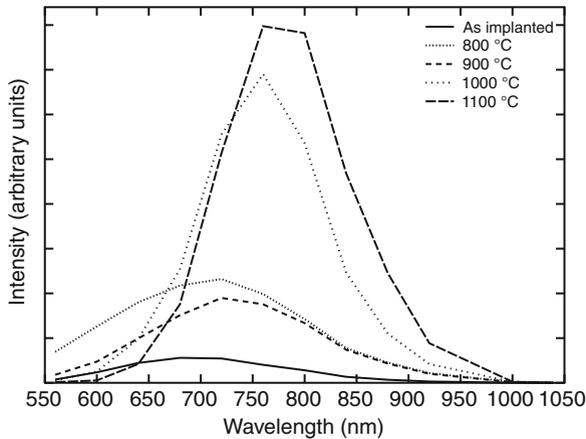
Fig. 4 Luminescence of Si nanoclusters in SiO_2 as a function of the annealing temperature. The results illustrate typical behaviour observed during nanocluster growth by Ostwald ripening: the intensity of the light emission increases, and the peak of the luminescence shifts towards longer wavelengths, when the nanoclusters grow in size for higher annealing temperatures [10]

We can summarize that the implantation energy defines the depth range of the atom and vacancy clusters, and that the implanted fluence defines the areal density of the clusters.

However, often an additional effect is present which can reduce the size variation. Since smaller clusters have a higher surface curvature and hence higher surface energy compared to larger ones, the smaller clusters are less stable. Hence if they can thermally emit atoms, the smaller clusters tend to shrink at the expense of the larger ones. Thus the smallest clusters will vanish and both the size and depth distributions of the clusters will become narrower. This process, known as Ostwald ripening [14], is well known in, e.g., emulsion chemistry and geology. By ending the high-temperature annealing at a suitable time, one can thus achieve a desired average cluster size with a fairly narrow size distribution.

Sometimes the size distribution can be narrowed more by utilizing a further ion beam treatment of the clusters after they have formed. The ion beams can cause erosion of atoms from the nanoclusters, thus reducing them in size. Since the largest clusters have a bigger cross-section area relative to the incoming ion beam, they are more likely to be reduced in size. Thus an "inverse" Ostwald ripening can be achieved, which in principle could even be used to obtain a monodisperse cluster size distribution. However, it is also possible that the emitted atoms start acting as seeds for new clusters, leading to the growth of smaller "satellite" clusters in the vicinity of the larger ones [15]. The clusters can by additional annealing or implantation also coalesce to larger clusters, also leading to different cluster size distribution.

The overall development of clusters in implantation and thermal annealing is illustrated in Fig. 3.

3 Formation of Nanovoids

Nanovoids, i.e., nanometer-sized empty regions of space inside a solid, are also of application interest because of the possibility of using them, e.g., as gettering sites for impurities in semiconductors, as seeds for cleavage, or because the inner surface formed could be active by itself. A conceptually obvious approach for forming a nanovoid could be to create defects in solids by ion irradiation and carry out annealing under conditions where more interstitials than vacancies migrate to the surface, leaving a vacancy excess in the solid. This concept, sometimes referred to as the "vacancy implanter" [16], has indeed been used to introduce vacancy clusters into solids. However, in many solids it is energetically favourable for vacancies to agglomerate into vacancy-type dislocation loops rather than voids. Hence this approach can seldom be used to directly form nanovoids, which are supposed to be roughly spherical in shape.

A more practical approach to form nanovoids is to use noble gas implantation to first form gas bubbles or crystals inside solids [17]. These bubbles form along the lines of processes described above. Thereafter if the lattice is heated above the temperature where the bubbles start to emit atoms, they will eventually be emptied of the gas atoms. Thus a region of roughly spherical voids is formed in the lattice in the same region where the gas bubbles were located originally.

4 Applications of Nanoclusters and Nanovoids

The nanoclusters and nanovoids formed in solids have potential for a wide range of practical applications. We describe here a few examples of possible applications.

4.1 Nanoclusters

The novel properties of nanophase particles are dominated by two major effects, namely (i) the increasing relative significance of the surface-energy contributions associated with the larger surface-to-volume ratio of small particles, and (ii) the unique characteristics of electrons in confined systems. The first effect largely determines the physical and thermodynamic properties of the nanoparticles (nps) or nanocomposites (e.g., melting points, solid phase transitions, bulk modules). Both the surface properties and electron confinement combine to produce novel electronic properties that can be manifested in a wide range of effects, such as a large nonlinear optical susceptibility, intense photoluminescence, altered band structures, and superparamagnetism.

Because of the large fraction of surface atoms, the first few atomic layers at the outside of the nanocluster will have other properties than the interior of the nanocluster (due to surface reconstruction, presence of defects, bending of electric bands, other charge carriers). One way to prevent surface effects is to passivate the nanocluster surface by embedding the nanoclusters in hosts with a large band gap

and similar structural properties. Ion implantation into ceramic oxides is a practical method to create embedded, electronically passivated semiconductor nanoclusters [18, 19]. An example of this is CdSe nanoclusters in MgO created by sequential Cd and Se ion implantation and subsequent thermal annealing [20].

Based on the observation that visible light can be obtained from Si-based materials, such as Si nanoclusters embedded in insulators like silica, silicon integrated photonics is anticipated to begin revolutionizing electronic devices, circuits, and systems [21]. Implantation of Si or Ge into silica or quartz leads to the formation of optically active nanocrystals in the silicon dioxide. Both Si and Ge nanocrystals have been demonstrated to have light emission in the optical and infrared light regions [22]. It is still not quite clear what the light emission mechanism is. Some of the light is clearly emitted from defects in the silica which are somehow stabilized by the nanoclusters [22, 23]. However, the emission often has a dependence on cluster size, indicating that quantum confinement in the crystals themselves also plays some role in the emission [24].

Even light amplification from Si nanocrystals in silica has been demonstrated, paving the way for a Si-based laser [25, 26]. For use in photonics, it is necessary to form patterns of nps in a way in which the sizes and locations of clusters can be controlled to form uniform patterns.

Nonvolatile flash memories based on nps are one of the promising routes to a further downscaling of CMOS technology. The increase of scale integration should involve new features such as Coulomb blockade, quantized charging effects, and single electron transfer [27, 28]. To appear at room temperature, these effects need nanometer-size nps in limited number and ideally a single particle. The problem is to fabricate these nps while controlling both their size, number, and position. It has been demonstrated [29] that a two-dimensional array of Si nps embedded into a thin SiO_2 layer (from 5 to 10 nm) can be synthesized by ultra-low energy ion implantation (1 keV) followed by thermal annealing. Synthesis of localized 2D layers of silicon nps embedded in a SiO_2 layer by a stencil-masked ultra-low energy ion implantation process has been proposed as an approach called "stencil-masked ion implantation process" to perform a spatially localized synthesis of a limited number of Si nps within a thin SiO_2 layer [30].

The Si nanocrystals in silica also have potential for use as a phase transition memory. Using laser processing, the crystal structure of the clusters can be altered [31], which could make for a memory device with extremely long storage times, since atomic phase transitions typically occur on much longer time scales than electron transport.

For the discussion of the synthesis of metal nanocluster composites in silica matrix, where the cluster concentration is below the percolation limit (dispersed clusters), we refer to a recent article by Mazzoldi et al. [32]. The catalyst nps prepared by ion implantation can be utilized for growth of carbon nanotubes via chemical vapor deposition [33]. Nickel ions were first implanted at room temperature into silicon dioxide. Postimplantation annealing was conducted to induce precipitation of implanted Ni atoms into nps. The samples were chemically etched to expose Ni nps on the surface. Finally, CNT growth on such prepared SiO_2 substrates was achieved

via chemical vapor deposition through decomposition of hydrocarbons. Nanoclusters with novel topologies, such as core-shell, hollow, multilayer, and triangular nanoclusters, have received attention because of their interesting morphologies, specific structures, and good optical, electric, or magnetic properties [34–37]. Meldrum et al. [38] observed the nanovoids in metal nanoclusters formed by ion implantation and subsequent annealing at a temperature higher than the corresponding bulk boiling points of the implanted elements. This phenomenon was attributed to the volume contraction during the solidification of the nanoclusters and the migration of implanted elements. Ren et al. observed that the nanovoids in nanoclusters formed by Cu ion implantation into silica evolve into different morphologies during subsequent annealing [39]. The same group has reported on controlling the morphology of Ag nanoclusters by high-dose Ag implantation into silica and subsequent annealing [40]. When the dose is increased over a certain value, the nanovoids gradually shrink and form a sandwiched nanocluster-nanovoid-nanocluster structure.

4.2 Nanovoids

Helium implantation followed by thermal annealing at elevated temperatures leads in a number of semiconductor and ceramic materials to void generation. This was first demonstrated for helium-implanted silicon [17, 41] and later on for helium and deuterium-implanted MgO [42, 43]. In cleaved MgO single crystals implanted with ^{3}He ions and subsequently thermally annealed, transmission electron microscopy (TEM) observations revealed the existence of sharply rectangular nanosize voids at a depth slightly shallower than the helium-implantation range [44]. The observation implies that nanocavities can be produced at a desired depth and with a desired size.

Mirabella et al. have studied the role of surface nanovoids on interstitial trapping in He implanted crystalline Si [45]. The difference with respect to the effects induced by the deep void layer was evidenced. B and He ion implantations were performed on Czochralski bulk and silicon-on-insulator samples in order to isolate the void surface region from the deep void layer. A remarkable reduction of B diffusion was recorded near half the projected range of He implantation, which leads to a boxlike shape in the distribution of mobile and electrically active B. Surface nanovoids (2–3 nm in size) caused the observed B diffusivity reduction through an enhanced recombination of self-interstitials.

The same group has made a systematic study on the formation of He ion implantation-induced nanovoids in Si and how they influence the self-interstitial supersaturation, thus affecting the diffusion and electrical activation of implanted boron in crystalline silicon [46]. The experiments demonstrate the role of nanovoids in reducing B diffusion at the first stages of postimplantation annealing. The effect was attributed to the interstitial trapping by the nanovoids that forces B to assume a boxlike profile. It was concluded that if the nanovoid density is high in the proximity of implanted B, implantation-related damage can annihilate at the internal dangling bonds of nanovoids, thus consuming the nanovoid layer.

References

1. B. Colombeau, N. E. B. Cowern, F. Cristiano, P. Calvo, N. Cherkashin, Y. Lamrani, and A. Claverie, Appl. Phys. Lett. **83**, 1953 (2003).
2. K. Nordlund, Comput. Mater. Sci. **3**, 448 (1995).
3. G. Nicotra, G. Franzo, and C. Spinella, Nucl. Instrum. Methods Phys. Res. B **257**, 104 (2007).
4. P. Haasen, *Physical Metallurgy* (Cambridge University Press, Cambridge, UK, 1978).
5. A. Luukkainen, J. Keinonen, and M. Erola, Phys. Rev. B **32**, 4814 (1985).
6. J. Keinonen, V. Karttunen, J. Räisänen, F.-J. Bergmeister, A. Luukkainen, and P. Tikkanen, Phys. Rev. B **34**, 8981 (1986).
7. J. Keinonen and V. Karttunen, Phys. Rev. B **37**, 8440 (1988).
8. R. C. Birtcher, S. E. Donnelly, M. Song, K. Furuya, K. Mitsuishi, and C. W. Allen, Phys. Rev. Lett. **83**, 1617 (1999).
9. K. O. E. Henriksson, K. Nordlund, A. Krasheninnikov, and J. Keinonen, Appl. Phys. Lett. **87**, 163113 (2005).
10. K. Aitola, T. Ahlgren, J. Keinonen, and L. Khriachtchev, 2009, to be published.
11. L. M. Brown and G. R. Woolhouse, Phil. Mag. **21**, 329 (1970).
12. H. Trinkaus, Radiat. Effects **78**, 189 (1983).
13. K. O. E. Henriksson, K. Nordlund, and J. Keinonen, Nucl. Instrum. Methods Phys. Res. B. **244**, 377 (2005).
14. W. Ostwald, Z. Phys. Chem. **22**, 289 (1897).
15. K.-H. Heinig, B. Schmidt, M. Strobel, and H. Bernas, in *X*, edited by X (MRS, ADDRESS, 2001), Mater. Res. Soc. Symp. Proc. **650**, 96 (2001).
16. R. Kalyanaraman, T. E. Haynes, M. Yoon, B. C. Larson, D. C. Jacobson, H. J. Gossmann, and C. S. Rafferty, Nucl. Instrum. Methods Phys. Res. B **175**, 182 (2001).
17. C. C. Griffioen, J. H. Evans, P. de Jong, and A. van Veen, Nucl. Instrum. Methods Phys. Res. B **27**, 417 (1987).
18. C. W. White, J. D. Budai, S. P. Withrow, J. G. Zhu, E. Sonder, R. Zuhr, A. Meldrum, D. M. Hembree, Jr., D. O. Henderson, and S. Prawer, Nucl. Instrum. Methods Phys. Res. B **141**, 228 (1998).
19. C. W. White, A. Meldrum, J. D. Budai, S. P. Withrow, E. Sonder, R. A. Zuhr, D. M. Hembree, Jr., M. Wu, and D. O. Henderson, Nucl. Instrum. Methods Phys. Res. B **148**, 991 (1999).
20. M. A. van Huis, A. van Veen, H. Schut, S. W. H. Eijt, B. J. Kooi, and J. T. M. De Hosson, Nucl. Instrum. Methods Phys. Res. B **216**, 126 (2004).
21. B. Jalali, Sci. Am. **296**, 46 (2007).
22. R. Salh, L. Fitting, E. V. Kolesnikova, A. A. Sitnikova, M. V. Zamoryanskaya, B. S. B, and H. J. Fitting, Semiconductors (2007).
23. K. Lieb and J. Keinonen, Contemp. Phys. **47**, 305 (2006).
24. L. Ding, T. P. Chen, Y. Liu, M. Yang, J. I. Wong, Y. C. Liu, A. D. Trigg, F. R. Zhu, M. C. Tan, and S. Fung, J. Appl. Phys. **101**, 103525 (2007).
25. L. Pavesi, L. D. Negro, C. Mazzoleni, G. Franzo, and F. P. F, Nature **408**, 440 (2000).
26. L. Khriachtchev, M. Räsänen, S. Novikov, and J. Sinkkonen, Appl. Phys. Lett. **79**, 1249 (2001).
27. S. Tiwari, F. Rana, H. I. Hanafi, A. Hartstein, E. F. Crabb, and K. Chan, Appl. Phys. Lett. **68**, 1377 (1996).
28. E. Kapetanakis, P. Normand, D. Tsoukalas, and K. Beltsios, Appl. Phys. Lett. **80**, 2794 (2002).
29. C. Bonafos, M. Carrada, N. Cherkashin, H. Coffin, D. Chassaing, G. B. Assayag, A. Claverie, T. Müller, K. H. Heinig, M. Perego, M. Fanciulli, P. Dimitrakis, and P. Normand, J. Appl. Phys. **95**, 5696 (2004).
30. C. Dumas, J. Grisolia, L. Ressier, A. Arbouet, V. Paillard, G. B. Assayag, A. Claverie, M. A. F. van den Boogaart, and J. Brugger, Phys. Stat. Sol. (a) **204**, 487 (2007).
31. L. Khriachtchev, M. Räsänen, and S. Novikov, Appl. Phys. Lett. **88**, 013102 (2006).
32. P. Mazzoldi and G. Mattei, Phys. Stat. Sol. (a) **204**, 621 (2007).

33. A. R. Adhikari, M. B. Huang, D. Wu, K. Dovidenko, B. Q. Wei, R. Vajtai, and P. Ajayan, Appl. Phys. Lett **86**, 053104 (2005).
34. X. Gao, J. S. Zhang, and L. Zhang, Adv. Mater. **14**, 290 (2002).
35. G. Schneider and G. Decher, Nano Lett. **4**, 1833 (2004).
36. E. Prodan, C. Radloff, N. J. Halas, and P. Nordlander, Science **302**, 419 (2003).
37. R. C. Jin, Y. C. Cao, E. C. Hao, G. S. Metraux, G. C. Schatz, and C. A. Mirkin, Nature **425**, 487 (2003).
38. A. Meldrum, J. Mater. Res. **16**, 2670 (2001).
39. F. Ren, C. Z. Jiang, C. Liu, and J. B. Wang, Appl. Phys. Lett. **88**, 183114 (2006).
40. F. Ren, C. Z. Jiang, C. Liu, J. B. Wang, and T. Oku, Phys. Rev. Lett. **97**, 165501 (2006).
41. S. M. Myers, D. M. Bishop, D. M. Follstaedt, H. J. Stein, and W. R. Wampler, Mater. Res. Soc. Symp. Proc. **238**, 549 (1996).
42. H. Schut, A. van Veen, F. Labohm, A. V. Fedorov, E. A. C. Neeft, and R. J. M. Konings, Nucl. Instrum. Methods Phys. Res. B **147**, 212 (1999).
43. A. van Veen, H. Schut, A. V. Fedorov, F. Labohm, E. A. C. Neeft, and R. J. M. Konings, Nucl. Instrum. Methods Phys. Res. B **148**, 768 (1999).
44. B. J. Kooi, A. van Veen, J. T. M. D. Hossona, H. Schut, A. V. Fedorov, and F. Labohm, Appl. Phys. Lett. **76**, 1110 (2000).
45. S. Mirabella, E. Bruno, F. Priolo, F. Giannazzo, C. Bongiorno, V. Raineri, E. Napolitani, and A. Carnera, **88**, 191910 (2006).
46. E. Bruno, S. Mirabella, F. Priolo, E. Napolitani, C. Bongiorno, and V. Raineri, J. Appl. Phys. **101**, 023515 (2007).

Plasma Etching and Integration with Nanoprocessing

Antti Nuottajärvi and Tarmo Suppula

1 Introduction

This chapter introduces plasma etching—an extensive and perhaps the most widely used micro- and nanoprocessing method both in industry and in research and development laboratories worldwide. The emphasis is on the practical methods in plasma etching and reactive ion etching when used for submicron and nanoscale device fabrication. The principles of plasma etching and reactive ion etching equipment for sample fabrication will be introduced.

2 Etching

"*Etching*" is a widely used expression for removing material from a surface, especially in connection with micromachining. Together with lithography, etching can be used for pattern transfer to produce a 3D-strucuture where a certain area on the surface is exposed to etching processes.

Etching methods in use are the wet etching and the dry etching processes. The conventional wet etching technique is widely used for its very good selectivity and ease of use for large scale patterns and structures. However, submicron pattern transfer using wet etching techniques becomes difficult, or even impossible, because of the nature of the process itself. Etch solutions exhibit surface tension and, in the case of submicron openings in resist or mask layers, the solution can bridge the submicron spaces, thus preventing further etching of the surface in those areas. Another problem arises from the capillary characteristic of the etch solvent. The solvent can delaminate masking films, causing the etch profile to be nonadjustable, and furthermore it can even etch metallic structures protected by masking films. Normally wet etching processes are very temperature critical, giving temperature-time dependency

A. Nuottajärvi (✉)
University of Jyväskylä, Nanoscience Center, University of Jyväskylä, Jyväskylä FIN-40014, Finland
e-mail: antti.nuottajarvi@jyu.fi

of the resulting structures. One other critical variable is the solution concentration. Variations in these parameters give rise to considerable variation in reproducibility.

As a solution to the above mentioned problems with wet etching, plasma enhanced dry etching processes have been developed, in particular for processing sub μm and nm scale structures. In the dry etching processes [1], the etching reaction is accomplished by chemical reactions with chemically reactive gases instead of liquids. In these reactions the atoms or ions of the etchant react with the substrate to form volatile products that are removed, e.g., by pumping from the etched areas.

Dry etching can be divided into two operational principles: physical ion etching and reactive chemical etching. Reactive ion etching (RIE) is a combination of these two. Chemical etching alone usually leads to an isotropic etching profile, while physical etching gives an anisotropic profile in the etched material. A combination of these two methods gives a profile intermediate between these two cases [2].

3 Etching Profile

Anisotropic etching profiles result from *physical etching*, i.e., bombarding the substrate with highly energetic chemically inert species or ions which collide with atoms of the material as they hit the surface of the material. This method is very directional, and when a suitable mask is used, the material is removed in the ion beam direction only and the mask pattern is transferred exactly onto the substrate. This is a most beneficial feature of dry etching when a high spatial frequency of submicron or nanometer features detail is required on the substrate. Anisotropic etching is well suited for etching silicon nitride (SiN), and when etching openings are required through one or more overlapping layers on the substrate. One disadvantage of physical etching only as an etching method is that areas close to the etching profile, i.e., the masking layer, are in danger of suffering physical erosion damage because the method is very nonselective.

When *chemical etching* plays the dominant role in the etching process, the resulting profile is isotropic. Here the energy of the ions hitting the etched surface is very small compared to the energies in physical etching, thus resulting in a very selective reaction in which only material intended to be removed is etched away while other materials, e.g., the masking layer, remain intact. When etching an opening through one material layer on top of another it is a matter of selectivity to obtain an etching that removes only the layer intended but not the material underneath (Fig. 1).

In many cases an intermediate form of the two dry etching methods described above is more convenient. *Reactive ion etching* [3], which in some cases is referred

Fig. 1 Isotropic and anisotropic etch profiles

to as reactive sputter etching, is a combination of both physical and chemical etching and the resulting etch profile is strongly dependent on the ratio of these two processes. In reactive ion etching, the surface to be etched is sputtered by ions of a chemically reactive material resulting in an etch profile that is intermediate between isotropic and anisotropic. Bombarding the surface with highly energetic, chemically reactive ions selective to the material produces volatile byproducts that are removed from the process through the vacuum system. The desired etching profile can then be obtained by adjusting the reaction conditions, i.e., pressure and acceleration voltage (of the ions). As a rough guideline, high pressures and low voltage result in an isotropic etching profile, whilst low pressure and high voltage gives an anisotropic profile.

Since plasma etching and physical sputtering are usually carried out in the same equipment but in a different process, the appropriate process parameters must be selected. Plasma etching is usually carried out at a pressure of a few tens of pascals. The plasma produces energetic neutral free radicals which react with the surface exposed to the process and attack it isotropically, resulting in an isotropic etch profile. For sputtering, on the other hand, a lower, down to a few mPa, process pressure is used and a suitable gas is injected. In the pure sputtering process a bombardment of highly energetic argon ions are used to eject the surface atoms from the surface (Box 2) [3–5] (Fig. 2).
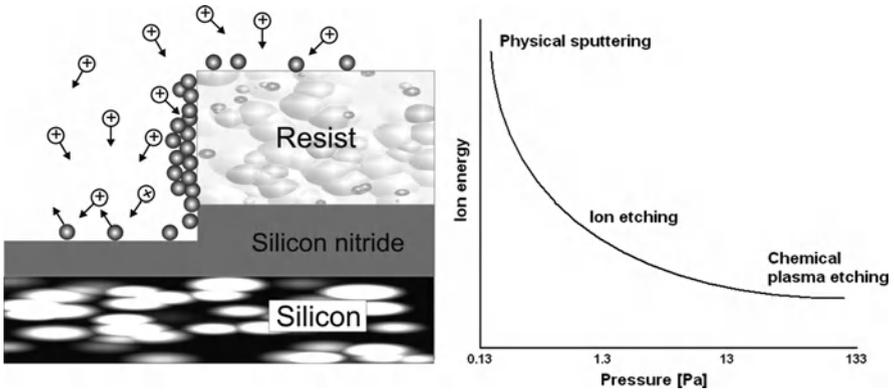
Fig. 2 Plasma etch process principle

4 Glow Discharges and Plasmas

"*Plasma*" is very widely used term. Here we consider a plasma to be a gas that contains electrically charged and neutral particles: electrons, positive and negative ions, as well as neutral atoms. The term "*glow discharge*" is used to describe a self-sustaining type of plasma used for processing. In a direct current glow discharge (DC exited plasma) a constant electric field is applied between two electrodes

under low pressure conditions where pressure can be adjusted between 0.13 Pa and 1300 Pa. When electrical potential of 100–500 V is applied between the electrodes, electrons accelerated in this field gain enough energy and ionise molecules of the gas between the electrodes. The positive ions are attracted towards the cathode by the electrical potential while electrons are attracted to the anode. These secondary ions and electrons subsequently collide with other atoms and ionize them. As long as the electrical potential is maintained, the cathode is being sputtered by impact of the positive ions generated in the plasma. A problem with DC sputter etching is that if the substrate is not conductive enough, it will charge-up by bombardment of ions. In radio frequency glow discharge the substrate charging problem is avoided, since the plasma is exited by an RF electric field that is applied on one electrode, while the other electrode is grounded (e.g., as part of the vacuum chamber around the plasma). Electrons that gain energy from the RF field collide with gas molecules resulting in ionization and dissociation. Current is carried in the plasma by both the equal number of electron and ion concentrations in the plasma. The accelerated electrons have higher mobility compared to the ions, so they reach the electrodes and the etching apparatus chamber walls easier. For zero net current flow, a DC blocking capacitor is needed between the electrode and the RF source. A negative charge is built up on the chamber walls and electrodes. This negative charge repels electrons and attracts positive ions. Under steady-state conditions, the flow of electrons and ions hitting the surface are equal and the glowing part of the discharge constitutes the most positive potential in the system [2, 6, 7].

Between the plasma and the electrodes, and the process chamber walls, there is a space, called the *plasma sheath*, over which the potential drops by a few hundreds of volts within a distance of a few millimeters. In most systems the grounded area, i.e., one electrode and the etching chamber walls, is larger than the electrode powered by RF. Because of that, the potential drop and the energy of ions hitting the RF powered electrode surface are higher than for the grounded electrode and the walls. The greater the difference between electrode areas the greater is the ratio of voltage drop across the sheaths. A substrate placed on a grounded electrode will then be bombarded by ions with low energy, and this case is generally termed *plasma etching*. On the other hand, a substrate placed on the RF powered electrode will undergo ion etching when the pressure condition inside the etching chamber is low, e.g., under 10 Pa. In this case the process is called *reactive ion etching (RIE)* [2, 6].

The plasma electron density is an important parameter. The electron density in the plasma normally lies in the range 10^9–$10^{12}\,cm^{-3}$ [2]. The electron energy is typically a few electron volts, which corresponds to approximately ten times higher temperature than the gas, which is normally well below 1000 K [2]. The relative concentrations of species in low-density etching glow discharges and high density etching glow discharges have different ratios of neutral radicals and charged particles (ions and electrons). In the glow discharge, electrons with high energy impact with the etching gas molecules and neutral radicals, creating ions and excited molecules. Neutral radicals have no electrical charge so they are not accelerated in the electric field, but they diffuse to the target area from all directions. Neutral radicals with
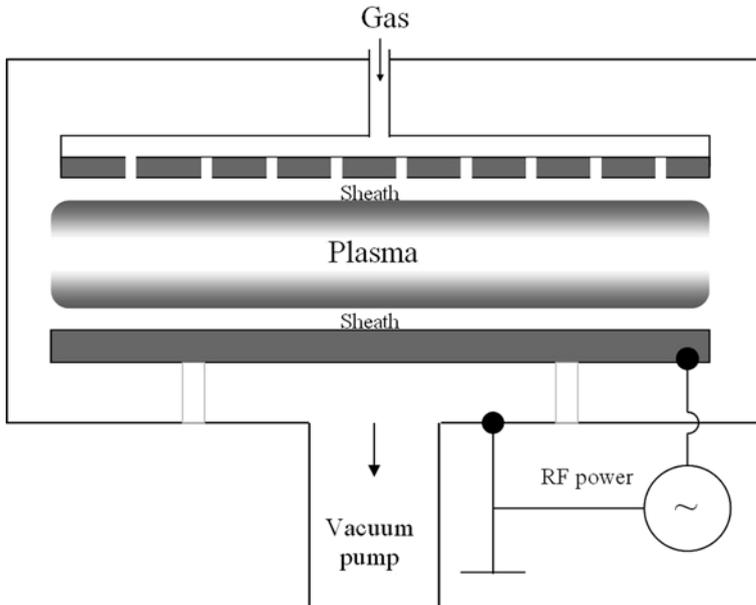
Fig. 3 Typical RIE setup

high chemical reactivity are typically created only with high electron temperatures. The plasma temperature remains normally at, or slightly above, room temperature, which allows thermally labile materials, e.g., polymers, to be plasma treated (Fig. 3).

When an ion in the glow discharge reaches the region close to an electrode it undergoes acceleration in the electric field and impinges on the electrode or substrate. In RIE these ions do not take part in chemical reactions at the surface but they do cause surface damage favorable for chemical reactions resulting in better removal of process byproducts.

The plasma potential governs the energy with which the ions bombard the surfaces. On the substrate, the ion energy gives etching directionality, while on the other surfaces, i.e., the etching chamber walls and electrodes, high energy ions may result in undesirable sputtering of contaminant particles. The plasma potential may be estimated by comparing the sheath thicknesses between the powered and grounded electrodes or by measuring the powered electrode voltages. Sheath thicknesses that are equal, or close to equal, correspond to a relatively high plasma potential. When RF voltage measured from the powered electrode is greater than DC voltage, it results in high plasma potential, and when these two voltages are close to equal, it then results in low plasma potential. The frequency of the RF affects the ion energy distribution. The most common RF employed is 13.56 MHz, which corresponds to approximately 22 eV in ion energy with respect to ground. Decreasing the frequency broadens the ion energy distribution.

5 Etching Rate and Selectivity, Loading Effect

The etch rate and selectivity are both highly dependent on the properties of the process itself. The chemical etch rate is generally exponentially dependent on the ratio of $E_a/k_B\,T$, where E_a is material and etch gas dependent activation energy of the chemical process, k_B is the Boltzmann constant, and T is the absolute temperature. In dry etch processes a high etch rate over that of the masking layer is desired.

Selectivity depends on the etch rates of the masking layer and the substrate material. The ratio depends on the difference in their activation energies, and it is strongly temperature dependent. However, when reactive ion etching dominates, the temperature dependence becomes negligible and the etch rate and the selectivity are more dependent on the ion energy or the ion flux. In certain processes these parameters have threshold values below which the etch rate becomes zero. Since the ion energy and the flux depend on the RF excitation power and frequency, the potential of the electrodes, and even the etching chamber geometry, these parameters may be adjusted to tune the selectivity by plasma properties [8].

Since the selectivity is dependent on the ion bombardment energy, the ion energy distribution (IED), or ion energy distribution function (IEDF), plays a governing role in plasma assisted processes. The IEDF is the fraction of ions hitting the substrate with energies E and $E + \Delta E$, and is determined by the ion motion from its creation to the collision with the substrate. In plasma etching processes, the IEDF is broad and the amplitude is close to zero because the ions undergo a number of collisions in sheath areas. In a RIE system, the ion energy is determined by the sheath potential, and the IEDF is broad because of energy consuming collisions with neutrals. For high density plasma systems, these have collision-less sheath areas and the IEDF can be narrow. In plasmas ionization is distributed over a volume, so the ions hit the substrate with different energies depending on the location where the ion was created. In other words, in low pressure processes the majority of ions traverse the sheath without undergoing collisions, resulting in a narrow peak; while in high pressure processing the collisions in the sheath area result in lower ion energies and IEDF broadening [6].

During the etching process the active species are depleted. This effect is called loading, and it strongly depends on the etched area. A basic rule is that the etching rate decreases when the area to be etched increases. It has been shown that in the case of SiN etched with CF_4 plasma, the etching rate decreases when the area exceeds $1000\,nm^2$ [2]. Loading effects can even cause nonuniform etching, or in the worst case, stop the etch process totally, and polymerization occurs. This can be a problem when multilayer structures are etched, because at the etched film–substrate interface the loading effect can strongly increase the etch rate, and the etched area diminishes rapidly. This is a problem especially when etching is used for making a film thinner, i.e., when etching the metallic strips to reduce the cross-sectional area, making them more electrically resistant. Loading can generally be avoided by maintaining the etched area as small as possible and the plasma volume as large as possible.

6 Fluorine and Chlorine Chemistry

In real situations when a dry etch process is used for device fabrication in a research laboratory or in a factory scale facility, one needs to consider carefully the etch process conditions, since there might be various materials in the chamber at the same potential as the substrate, not all of which are desirable to be etched.

The chemistry between the gas or gases used and the etched materials is somewhat complex, but some basics are discussed in the following.

In dry etching and RIE, chlorine- and fluorine-based processes are used.

6.1 Etching Silicon, Silicon Dioxide, and Silicon Nitride

The gases used for etching Si, SiO_2, and SiN are usually fluorocarbons like CF_4. The etch rate of fluorine atoms against Si is given by [6]:

$$R_{F(Si)} = 2.91 \times 10^{-12} \sqrt{T} n_F \exp(-E_a/kT) \quad [\text{nm/min}]. \tag{1}$$

In the case of CF_4, the reaction with plasma electrons produces trifluoromethane and fluorine:

$$e^- + CF_4 \leftrightarrow CF_3 + F + e^-. \tag{2}$$

The F atoms react with the silicon film to produce SiF_2 and SiF_4.

Other commonly used fluorine forming plasmas that are used for etching silicon are a combination of CF_4 and O_2, SF_6, SF_6/O_2, SiF_4, SiF_4/O_2, NF_3, etc. [9–12].

When oxygen gas is included in the process, the yield of fluorine atoms is increased. The oxygen reacts with CF_x radicals, forming CO, CO_2, COF_2, and F. This also leads to less polymerization, since CF_x is depleted in the reaction. If CF_4 is diluted too much with O_2, the etch rate decreases. Adding oxygen in F-plasmas has proved to be an effective way to tune Si/SiO_2 etch selectivity. The fluorine etch processes are usually preferred to isotropic silicon etching when there is no need for pattern transfer, but blanket etching to, e.g., form metallic free hanging isolated bridge-like structures (Fig. 4).

Along with the fluorocarbons, plasmas producing chlorine atoms are also used for etching silicon. The anisotropy can be well controlled in chlorine-based etching processes. A useful feature is that chlorine atoms can etch doped silicon chemically even without ion bombardment, while in the case of undoped silicon, ion bombardment is needed. Chlorine molecules usually do not etch silicon even if they adsorb on the surface.

The feedstock gases used in chlorine-based etching are usually Cl_2, CCl_4, CF_2Cl_2, and CF_3Cl [10]. Gas mixtures like Cl_2/C_2F_6, Cl_2/CCl_4, and C_2F_6/CF_3Cl are also used when a particular etch rate is required. Sometimes gases are mixed with argon gas to accelerate the etching rate. Gas mixtures with O_2 lead to very clear anisotropic
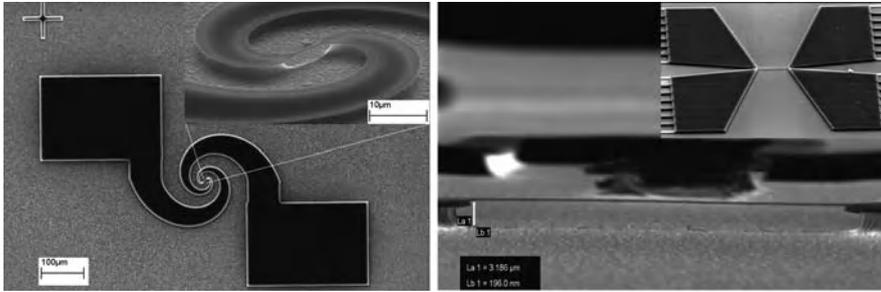
Fig. 4 Free hanging structures processed in University of Jyväskylä Nanoscience Center [18]

etch profiles because of a phenomenon called sidewall passivation, where an oxide layer is formed on the side walls while etching continues vertically. However, if O atoms are significantly present, they can cause problems by forming microscopic etch-resistant oxide areas on the surface. This phenomenon is usually called grassing [13], or formation of black silicon. This is because the surface becomes very rough, since there are areas that are not etched properly. In most cases this unwanted phenomenon originates from water vapour that can be released from films during the etch process or that can be introduced if the gas source has an appreciable water content.

Silicon etch plasmas can also be used for etching silicon dioxide, SiO_2, and silicon nitride, SiN. When etching SiO_2 a fluorine-based process is usually required, since etching with chlorine is thermodynamically an endothermic reaction. Conventionally, the gases used for SiO_2 etching are CF_4 and CHF_3. CF_4 mixtures with O_2 or with H_2 are used to control selectivity. For selectivity between Si and resist polymers, a carbon-containing and fluorine-deficient plasma can be achieved by using CHF_3, CF_4/H_2, C_2F_6, or mixtures of these gases. If selectivity for Si is not important, the gases CF_4 and CF_4/O_2 are widely used. Carbon containing gases such as CO, CO_2, and COF_2 thus enhance oxide etching. Gases forming fluorine atoms, like SF_6 and NF_3, are also useful for oxide etching. In etching SiO_2, the anisotropy is good because the spontaneous reaction of SiO_2 with F is slow.

Etching of silicon nitride is often referred to as "a not very well understood" process. This originates from the question of the identity of etch products from SiN etching. As it was with silicon and CF_4:

$$e^- + CF_4 \leftrightarrow CF_3 + F + e^-, \tag{3}$$
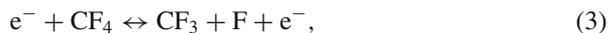
the fluorine free radicals react with SiN, or Si_3N_4, to be more exact:

$$12F + Si_3N_4 \rightarrow 3SiF_4 + 2N_2. \tag{4}$$

The reaction product SiF_4 is easily removed from the process chamber by pumping. Adding oxygen to the process enhances the number of fluorine atoms in the

plasma, giving too little polymerization, which inhibits the etch process. Inclusion of N, H (and F) in the process also allows formation of passivating layers.

SiN in etching behaviour is intermediate between Si and SiO_2. When a fluorine chemistry is used in an etch process with no ion bombardment, the etch rate is highest on Si and lowest on SiO_2, with SiN in between. In the case of chlorine chemistry and ion bombardment, the etch rates follow the same rule. Typically the RIE rates of Si, SiO_2, and SiN are of the order of 300, 60–100, and 100 nm/min, respectively. The rate depends on the electrode potential but is rather independent of temperature, indicating that ion bombardment is important in the processes [6, 9, 14, 15].

The selectivity of silicon nitride to silicon and silicon oxide can be tuned by selection of feedstock gases. For SiN/Si fluorine-deficient fluorocarbon, plasmas like CHF_3 and $CF_4 + H_2$ are the best choice. In these plasma processes the carbon blocking is selectively present on the Si but removed from SiN. For SiO_2/SiN fluorine-deficient fluorocarbon, plasmas with no hydrogen are the best choice. Those plasma sources are CF_4/C_2F_4, C_2F_6, etc. Here also carbon blocking is present on Si, as in SiO_2/Si selective etching. Fluorine-rich plasmas with low self-bias voltages, like the one with SF_6, is preferred for SiN/SiO_2 selectivity. In the case of SiN/SiO_2, chlorine chemistry can also be used [16].

In lithographic pattern transfer using a polymer resist mask, the resist must be able to withstand both ion bombardment and the reactive species of the plasma. However, RIE selectivity is relatively low (about 3) for SiN, compared to a regular photo-resist for CF_4/O_2 (4%) [17]. Pure oxygen plasmas are used to remove organic residuals and can also be used for selective stripping of the whole resist layer from the substrate (plasma ashing).

6.2 Etching Metals

Most metals react readily with oxygen and water vapor. Hence, the etch processes on them are more sensitive than on, e.g., silicon. Oxygen and water vapor must be excluded from the process by long pumping or scavenging. Normally ion bombardment is required to remove the natural oxide from the metal because the oxygen-metal bond is often very strong. Both fluorocarbons and chlorocarbons are widely used for metal etching. Both are suitable for oxide removal. It should be borne in mind that halocarbon vapors can cause polymerization on the surface and thus interfere with the metal etching process.

Difficult metals or materials to etch are those that do not form volatile products or for which temperature becomes a problem. Fe, Ni, and Co have no volatile halides and carbonyl formation is difficult. Au, Ag, and Pt have a problem of sputtering, together with no volatile halides, although gold is etched successfully with $CClF_3$ [6]. Copper forms volatile products only above 473 K. Aluminum oxide was already mentioned, as its reaction with chlorine is thermodynamically an exothermic reaction.

7 Plasma Etching Systems

Parallel plate reactor systems are widely used for many etching processes and can be considered to be a standard method for most applications. This reactor type can be used for plasma etching and RIE processing. Wafers or chips placed on the lower electrode are exposed to energetic radiation from the plasma generated between the two electrodes by means of RF coupled to one of the electrodes. In many applications (RIE) the lower electrode is connected to RF and it is smaller in area compared to the upper electrode, which, together with the etching chamber walls, is grounded. As mentioned above, the ratio of electrode areas determines the plasma potential distribution which governs the ion bombardment. In the case of physical etching, the upper electrode is connected to the RF power while the wafers are placed on the grounded lower electrode.

The process gases are fed to the system through the upper electrode. This often has a showerhead geometry, i.e., the electrode face area is covered with small holes that distribute the gases uniformly between the electrodes. The electrodes are connected to an RF generator by a so-called matching unit, that ensures optimum power transfer to the plasma. The vacuum is usually produced by a turbo pump. If the pre-vacuum is generated by a rotary pump with oil, special care needs to be taken. An inert Teflon oil should be used, because many etch byproducts are reactive. This is particularly important in all applications where oxygen is pumped through the system.

The process gases are fed into the system through separate mass-flow controllers for each gas used. In modern systems the chamber operating pressure is controlled by automatic feedback valves, while in older systems the process pressure and the vacuum were manually controlled. In modern systems the mass-flow controllers are placed away from the plasma etch tool. In a clean room environment it is useful to have a separate "gas pod" in the grey area, from which only one gas line feeds the instrument inside the clean area. PC-based controllers for the gas pod, instrument, and vacuum system provide a convenient way to maintain system parameters at set values. Along with the vacuum and gas flow rate, the PC often monitors parameters such as etch time, forward and reflected RF power, DC bias, process pressure, and electrode temperature. Usually a capacitance manometer is employed to measure the process pressure. In some systems electrode cooling with liquid nitrogen is possible. The helium-assisted lower electrode is used to secure a good thermal connection between the wafer and the electrode.

When operating a system it is important that the chamber is cleaned frequently. Especially in fluorine-based systems, process byproducts coat the chamber walls. To remove those unwanted materials from the system, mechanical cleaning and oxygen plasma cleaning are needed. If cleaning is not done on a frequent basis, these byproducts start to redeposit on the substrate being processed. This can prevent the etching and may be detrimental for post-processing steps. To detect this kind of contamination it is important to monitor the value of the DC bias. If this starts to decrease compared to the standard process, this is indicative of possible byproduct contamination. Another key parameter is the chamber base pressure. If the time to reach that pressure value increases, this is also indicative of possible contamination.

8 Etch Parameters and Recipes

The parameters used at JY/NSC [18] in etch recipes for real situations are presented in Table 1 for some common recipes at facilities with Plasmalab 80 RIE [19] and Axic Benchmark RIE tools. Some Ar sputter recipes and cryo-etch recipes as well as oxygen cleaning recipes are also included and discussed briefly.

Table 1 Recipe parameters used in some common processes at University of Jyväskylä Nanoscience Center [18]

Recipe	Etch rate [nm/min]	Gas-1	Gas-2	Time (s)	Temperature (K)	Pressure (Pa)	Power (W)
Alumina etch	18	SF_6	Ar	360	283	267	250
Ar sputter etch		Ar			303	20	250
Silicon nitride etch	50	CHF_3	O_2	120	303	733	150
O_2 clean chamber		O_2		120	303	533	200
O_2 flash (PMMA)		O_2		30	303	533	60
Oxide etch	42	CHF_3	Ar	300	293	40	200
Poly-Si etch		SF_6			303	08	60
Si:SiO etch		SF_6	O_2		303	120	50
Si cryo etch	540	SF_6	O_2	600	173	1267	200
Ti etch	10	SF_6	O_2	30	303	100	200
O_2 polymer strip etch	150	O_2		120	293	1293	100

One can see that the selective Si:SiO (300:1) etch gives a high selectivity rate close to room temperature. It can also be seen that some oxygen speeds up the etching. The so-called cryo-etch for silicon gives a 500 nm/min etch rate at 173 K. Liquid nitrogen is used for cooling the electrode, and a helium backing is used for thermal contacting. Some additional recipes for etching alumina and titanium as well as argon sputtering recipes are also given.

The oxygen polymer strip etch and oxygen cleaning recipes for removing organic residuals from PMMA after development have different powers and times, to prevent severe erosion of PMMA when only a surface layer of organic residuals must be removed. Chamber cleaning for contamination control is important for maintaining reliability of the recipies. Chamber cleaning is carried out by "scrubbing" the chamber mechanically with a suitable cloth and solvent, like high-class clean room tissue and acetone, followed by an etch.

8.1 Advanced Techniques

For high aspect ratio reactive ion etching some techniques like deep reactive ion etching (DRIE), including cryogenic DRIE, and the so-called Bosch process, have been developed. Both techniques result in a highly anisotropic etch profile with steep-sided trenches with aspect ratios exceeding several tens to one.

In cryoassisted RIE the substrate is cooled to slow down the chemical reactions on the sidewalls that result in an isotropic etch profile. Normally the substrate temperature is reduced to 165–175 K, as in the Si cryo etch recipe (Table 1). While isotropic etching is inhibited by cooling, the horizontal surfaces are bombarded by ions, which results in their being etched away with an improved etch rate [20, 21].

In two Bosch processes, two etching modes are repetitively cycled. The first step is a regular isotropic etch process. Subsequently, a chemically inert passivation layer is deposited. The etching phases are carried out for a few seconds at a time, and between these steps the passivation layer is deposited all over the substrate to prevent chemical attacking and etching. The directional ion bombardment sputters away the passivating layer from the horizontal surface at the bottom of the trenches, exposing the substrate to the chemical etching reactions. These two steps are repeated cyclically. The cycle time determines the etch profile. Short cycle times give smoother walls, while longer cycle times result in a higher etch rate. The etch rate in the Bosch process is typically a few micrometers per minute.

Plasma etching and reactive ion etching offer a powerful tool for nano-scale machining of surfaces. Uses for these technologies can be found over a wide range of industries in the research and development field. Although semiconductor materials are the main focus of plasma etching, the technique can be used to etch various metals and other materials as well. In addition to the most common reactor type discussed here, several different reactor types have been designed. Recipe libraries supplied by instrument manufacturers and research groups are a good starting point when a new etching process is being developed. Fine-tuning all the parameters to yield successful process results, and good reproducibility in contrast to time and money spent, often requires considerable effort. One wildcard that remains is that the instruments used are different, and a process or recipe that gives good results in one system does not work exactly the same way in another. In industries where certain processes are run constantly, dedicated instruments are used for a single recipe, while in R&D facilities, recipes and processes need to be modified and changed constantly.

References

1. H.W. Lehmann and R. Widmer, J. Vac. Sci. Technol. **15**(2), (1978).
2. A. Lindell, *Nanofabrication by atomic force microscopy, Electron beam lithography and reactive ion etching*, Department of Physics, University of Jyväskylä, Research Report 8/2000.
3. J.A. Bondur, J. Vac. Sci. Technol. **13**(5), (1976).
4. S. Somekh, J. Vac. Sci. Technol. **13**(5), (1976).
5. C.M. Melliar-Smith, J. Vac. Sci. Technol. **13**(5), (1976).
6. J.A. Mucha, D.W. Hess, E.S. Aydil, "Plasma Etching" in *Introduction to Microlithography*, second edition, American Chemical Society, Washington, (1994).
7. J.W. Coburn and H. F Winters, J. Vac. Sci. Technol. **15**(2), (1978).
8. W.M. Holber and J. Foster, J. Vac. Sci. Technol. A **8**(5), (1990).
9. M.J. Vasile and F.A. Stewie, J. Appl. Phys. **53**(5), (1982).
10. R. Legtenberg, H. Jansen, M. de Boer, and M. Elwenspoek, J. Electrochem. Soc. **142**(6), (1995).

11. B.E.E. Kastenmeier, P.J. Matsuo, J.J. Beulens, and G.S. Oehrlein, J. Vac. Sci. Technol. A **14**(5), (1996)
12. S. Gomez, R. Jun Belen, M. Kiehlbauch, E.S. Aydil, J. Vac. Sci. Technol. A **22**(3), (2004).
13. H. Jensen, M. de Boer, R. Legtenberg, M. Elwenspoek, J. Micromech. Microeng. **5**, (1995)
14. S. Matsuo, J. Vac. Sci. Technol. **17**(2), (1980).
15. B.E.E. Kastenmeier, P.J. Matsuo, and G.S. Oehrlein, J. Vac. Sci. Technol. A **17**(6), (1999)
16. S. Suto, N. Hayasaka, H. Okano, and Y. Horiike, J. Electrochem. Soc. **136**, 2032 (1989)
17. R.G. Poulsen, J. Vac. Sci. Technol. **14**(1), (1977).
18. University of Jyväskylä, Nanoscience Center, http://www.jyu.fi/science/muut_yksikot/nsc/en/
19. Plasmalab 80 Plus, Oxford Instruments Plasma Technology
20. F. Marty, L. Rousseau, B. Saadany, B. Mercier, O. Francais, Y. Mita, and T. Bourouina, Microelectron. J. **36**, (2005)
21. S. Tachi, K. Tsujimoto, and S. Okudaira, Appl. Phys. Lett. **52**(8), (1988).

Focused Ion Beam Machining and Deposition

Yongqi Fu and Lumin Wang

1 Introduction

The focused ion beam (FIB) technique was developed mainly during the late 1970s and the early 1980s. The first commercial instruments were introduced more than a decade ago [1]. Modern FIB systems are becoming widely available in semiconductor, microelectronics, and processing environments, as well as in failure analysis and chip-design centers. The technology enables localized milling and deposition of conductors and insulators with high precision, hence its success in device modification, mask repair, process control, and failure analysis [2–6]. Also, the preparation of specimens for transmission electron microscopy (TEM) and the trimming of thin-film magneto-resistive heads (for magnetic storage disks) are important applications of FIB [7, 8]. Only recently, a number of authors have reported the use of FIB in micromachining applications for microelectronic and mechanical systems [9, 10].

Focused ion beams have many advantages over other high-energy particle beams. For example, compared to photons or electrons, ions are much heavier and can deposit much greater energy densities on the target to directly write or mill patterns on hard materials, such as semiconductor, metal, or ceramic substrates. On the other hand, photons and electrons can effectively write on or expose only soft materials, such as photoresists (SU8, AZ1350, AZ-P4400, etc.) or e-beam resists (PMMA). The resists are normally used as the media to transfer the patterns from mask to silicon or other substrates. Commonly used techniques are photolithography, laser direct writing, and e-beam lithography (also called e-beam writing). Therefore, the direct writing capability of the FIB can reduce various difficulties and defects caused by the masks and resists in pattern transfer. Also, compared to e-beams, FIB does not significantly generate high-energy backscattered electrons, which often limit the minimum linewidth attainable by the e-beam lithographic or lift-off process. On the other hand, because ions are much heavier, the lateral scattering of the FIB is

Y. Fu (✉)
School of Physical Electronics, University of Electronic Science and Technology of China, Chengdu 610054, P.R. China
e-mail: yqfu@uestc.edu.cn

relatively low, resulting in the dose being well confined in the irradiated regions. This region is limited by the extent of the deposited energy; hence, a fine FIB can directly write a very narrow line in thin material.

1.1 Working Principles of FIB Machine

With the recent advent of bright sources such as the liquid metal ion source (LMIS) in the late 1970s and advances in ion optics in the late 1980s, FIB technology has been developing rapidly and so have its applications. The range of materials used in FIB systems is also expanding, further increasing the extent of their applications. The ion sources that are currently available include Al, As, Au, B, Be, Bi, Cs, Cu, Ga, Ge, Er, Fe, H, In, Li, Ni, P, Pb, Pd, Pr, Pt, Si, Sn, U, and Zn. Many of these ion species are produced from liquid-metal alloy sources because of the high melting temperature and the reactivity or volatility associated with the pure metal species. Ar, B, and P ions are particularly interesting because of their use in implantation of semiconductor materials. The most common ion species used for microfabrication are As, Be, Ga, and Si.

There are two types of FIB machines. One type is based on a single ion optical column, and the other is a so-called dual beam system with separate e-beam and ion beam columns integrated together in the vacuum chamber. The two beams may be tuned to focus on the same area of the sample, so the operator can use the e-beam (SEM) for imaging after FIB machining to avoid the damage caused if FIB imaging is carried out. A schematic diagram of an FIB ion column is shown in Fig. 1. The structure of the column is similar to that of a scanning electron microscope, the major difference being the use of an ion source producing a gallium ion (Ga^+) beam instead of an electron beam. A vacuum of about 1×10^{-5} Pa is maintained inside the column. The ion beam is generated from a LMIS by the application of a strong electric field. This electric field causes the emission of positively charged ions from a liquid gallium cone, which is formed on the tip of a tungsten needle. A typical extraction voltage is 7000 V. The extraction current under normal operating conditions is $\sim$2 μA [11]. An ion-optical configuration employing two lenses in a condenser-objective combination is widely used. One advantage of this approach is that, with a beam defining aperture in front of the second (objective) lens, the strength of the first lens can be varied to control the current through the system. The two-lens approach based upon a tetrode gun-condenser lens combination, with an einzel lens objective, has been used in lithography, microscopy, and micromachining applications. After a first refinement through the spray aperture, the ion beam is condensed in the first electrostatic lens. The upper octopole then adjusts the beam stigmatism. The ion beam energy is typically between 10 and 50 keV, with beam currents varying between 1 pA and 10 nA. Using the variable aperture mechanism, the beam current can be varied over four decades, allowing both a fine beam for high-resolution imaging on sensitive samples and a heavy beam for fast and rough milling. Typically seven preset values of beam current can be selected.
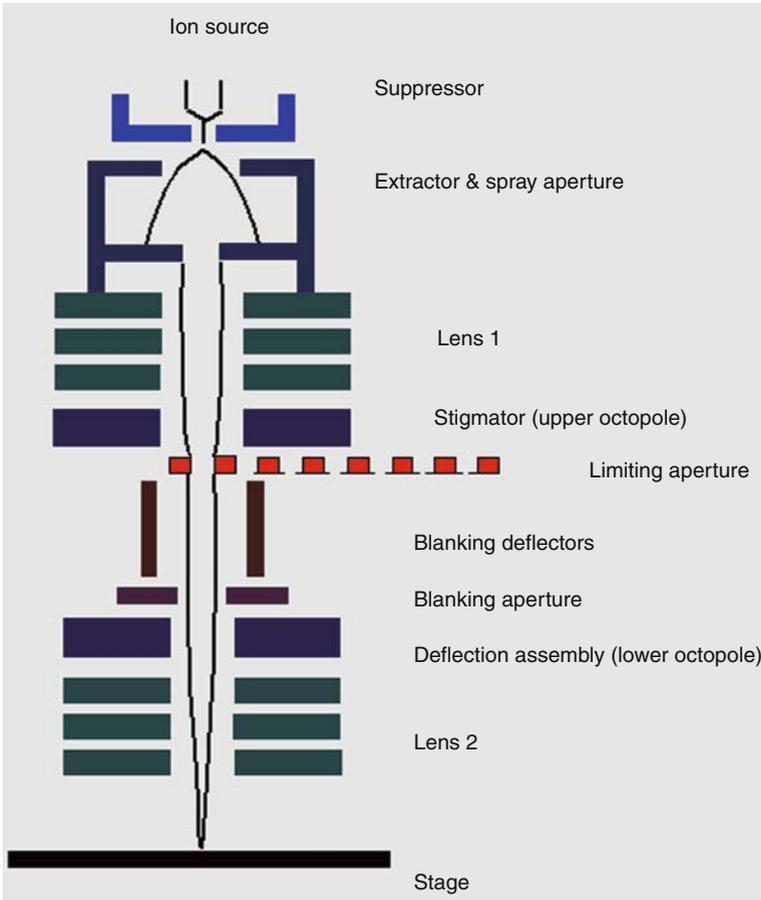
Fig. 1 Schematic configuration of FIB machine and its working principle

Blanking of the beam is accomplished by the blanking deflector and aperture (Fig. 1), while the lower octopole is used for raster scanning the beam over the sample in a user-defined pattern. In the second electrostatic lens, the beam is focused to a fine spot, enabling a best resolution in the range of sub 10 nm. The multichannel plate (MCP) is used to collect secondary particles for imaging.

Like all optical systems the ion-optical column has many types of aberrations in focusing the ion beam, e.g., spherical aberration, chromatic aberration, and astigmatism, etc. Therefore, the concepts of focusing spot size, depth of focus, and focal length, etc., are also applicable to the ion-optical column. Three main types of aberrations exist in the FIB ion column: spherical aberration d_s, chromatic aberration d_c, and astigmatism δx_c.

Spherical aberration produces a disc of least confusion with diameter strongly dependent upon the ion beam convergence half-angle α, and can be expressed as:

$$d_s = \frac{1}{2}\alpha^3 C_s, \tag{1}$$

where C_s is the spherical aberration coefficient of the lens referred to the image side. The diameter of the disc of least confusion, due to the energy spread $\Delta E_{1/2}$ of ions, having final energy E, is

$$d_c = \alpha \frac{\Delta E_{1/2}}{E} C_c \tag{2}$$

for the chromatic aberration coefficient C_c referred to the image side. $\Delta E_{1/2}$ refers to the full width and half maximum (FWHM) axial energy spread of ions from the LMIS, Gallium ion source for most commercial FIB machines.

Astigmatism is caused by the higher fractional energy spread obtained under these circumstances. It is the introduction of distortion into the beam by the nonuniform fringing fields at the entrance and exit of the analyzer. Some degree of astigmatism will also be introduced by the act of filtering, so that the beam of selected ions will be stretched asymmetrically due to their spread in energies.

$$\delta x_c = -\frac{ZB}{4}\left(\frac{2q}{M}\right)^{1/2}\frac{\delta V}{V^{3/2}} \tag{3}$$

where B is the magnetic field, q is the velocity of an ion of charge, M is the mass, V is the accelerating voltage, and Z is the mass selection size. In practice, this is sufficiently small to be removed by astigmators, located downstream of the filter. The depth of focus for the ion-optical column is expressed as

$$DF = \frac{2d_0 M}{NA}\sqrt{\frac{V_i}{V_0}} \tag{4}$$

where d_0 is the spot diameter in the focal plane, NA is the numerical aperture, M is the magnification, V_0 is the source extraction energy, and V_i is the final energy of the system (beam potentials).

The finite size of the source determines the ultimate limit on the probe size. If a lens with small aberration coefficient C_s and C_c is operated at vanishingly small values of α, so that the aberration discs vanish, the final probe will have a finite size determined solely by the magnification of the lens system and the diameter of the active ion emitting area of the LMIS tip d_t (see Fig. 2). Sudraud and coworkers reported that the diameter can be as small as $d_t \sim 10\,\text{nm}$ [12]. For an ion optical system of magnification M, the contribution this makes to the final probe diameter is

$$d_0 = Md_t \tag{5}$$

where the magnification M of the lens depends upon the ratio of voltages V_o and V_i in the image and object spaces, according to the Helmholtz-Lagrange relationship

Fig. 2 Gallium ion source in a typical FIB instrument

650 μm
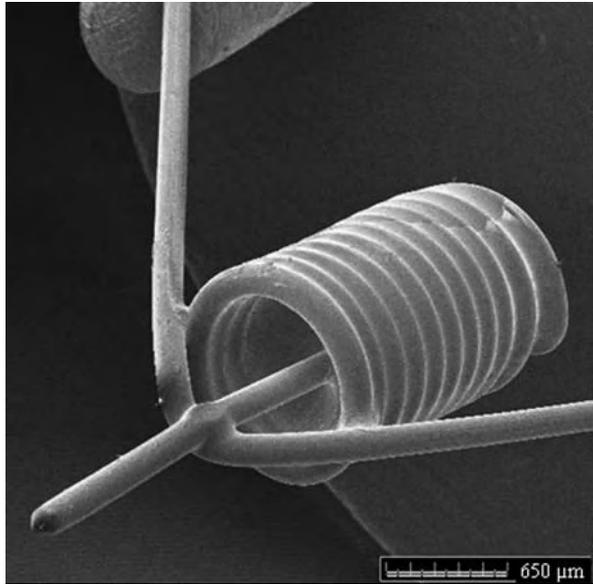
$$M = \frac{\alpha_0}{\alpha} \left(\frac{V_0}{V_i} \right)^{1/2} \tag{6}$$

where α_0 and α are the half convergence angles at incidence plane and image plane, respectively. The final spot size d is then determined by

$$d = (d_0^2 + d_c^2 + d_s^2)^{1/2} \tag{7}$$

Three distinct regions are predicted by above model:

(a) At high currents (large angular aperture), spherical aberration dominates and $d \propto I^{3/2}$.
(b) At intermediate values of current and α, chromatic aberration becomes the dominant effect, so that $d \propto I$.
(c) For the smallest value of beam current I, with the lens stopped down to very small values of α, the physical size of the source dominates and d is independent of I.

The focused ion beam energy distribution has in general a Gaussian profile, and is expressed as

$$J(x, y)/J_0 = e^{\frac{-f(x^2+y^2)}{r^2}} \tag{8}$$

where $J(x,y)$ is the beam intensity distribution at point (x, y); J_0 is the peak intensity at beam center; e^{-f} represents beam intensity insignificant to surface roughness; and r is the effective beam radius.

Secondary ions, secondary electrons, X-ray, γ-ray, and Auger electrons are generated when the ion beam hits on the substrate surface, as shown in Fig. 3. Both secondary ions and secondary electrons are collected by suitable detectors for imaging and metrology. The former is useful for imaging insulating materials, and the latter for metal and semiconductor materials imaging. The secondary electrons are also used for monitoring ion probing substrates with multilayers. This is called the end-point detection function and is commonly used in IC chip modification.
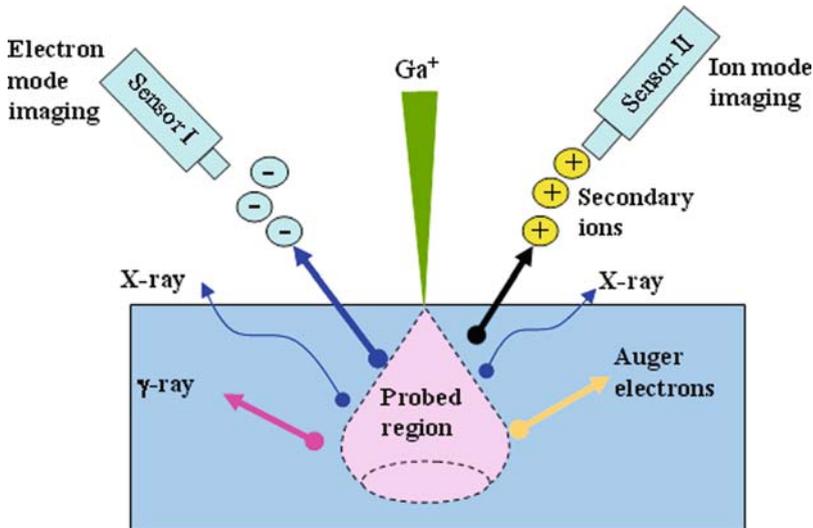
Fig. 3 Schematic diagram of ion sputtering on the substrate surface

1.2 Applications of FIB

FIB is used for repairing defects in optical and X-ray masks, circuit modification, device modification, failure analysis of circuits and device cross-sectioning, TEM sample preparation, microscopy, metrology (SIMS), direct resistless lithography, and direct maskless implantation of dopant into devices. With FIB it is possible on a sub μm or even nm scale to remove selectively a predetermined amount and kind of material from a selected volume of a target, and it is also possible to deposit at points of choice certain metals and insulators. Holes can be opened and refilled. By carefully removing volumes of material, one can use the FIB as a scalpel to carve out tiny, thin slices of material for examination in other instruments. Tiny wires can be fabricated to connect parts of ICs together that were left out of a design, or connections can be broken. If suitable ions of sufficient energy are used, it is possible

to implant directly into regions in devices which cannot be done in a reasonable manner using normal masked implantation. In addition to the above mentioned IC chip modification, TEM sample preparation, and semiconductor failure analysis, FIB has the following additional applications:

(1) metallurgical detection and failure analysis,
(2) three-dimensional (3D) micro/nano-fabrication,
(3) fabrication of micro-optical elements and plasmonic nanostructures,
(4) trimming of thin-film magnetoresistive heads for magnetic storage disks,
(5) imaging for biosamples, and
(6) micromachining for MEMS/NEMS.

2 FIB Direct Machining

There are two scan modes for the conventional commercial FIB machines: raster scan and serpentine scan, as shown in Fig. 4. Normally, the raster scan is used for milling 2D and 3D structures, while serpentine scans are used for cross-sectioning. The beam spots overlap each other during each scan step.
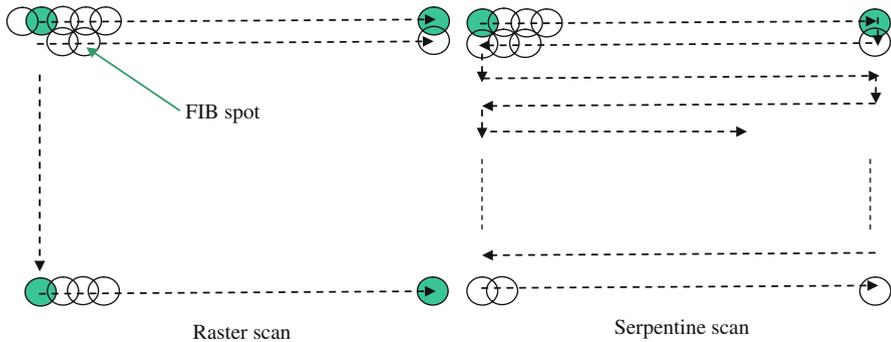
Fig. 4 Schematic diagram for raster scan and serpentine scan used in FIB machines

2.1 FIB Two-Dimensional Patterning

Two-dimensional (2D) FIB patterning can be realized by creating 2D patterns or artificial user defined bitmap patterns which are designed using standard professional drawing software such as AutoCAD, CorelDraw, Solidworks, ProEngineering, etc.

Using the above method and procedures, the structure of microgear was designed by AutoCAD, and milled by the FIB machine Micrion 9500EX. As can be seen from Fig. 5, the gear pattern was milled successfully. Although the shapes are complex and a large beam-spot size is used, the manufactured structure has a rough periphery

Fig. 5 SEM image of a
microgear milled using 2D
bitmap pattern method with
process parameters of ion
fluence corresponding to
$5\,nC/\mu m^2$ and aperture size
$150\,\mu m$
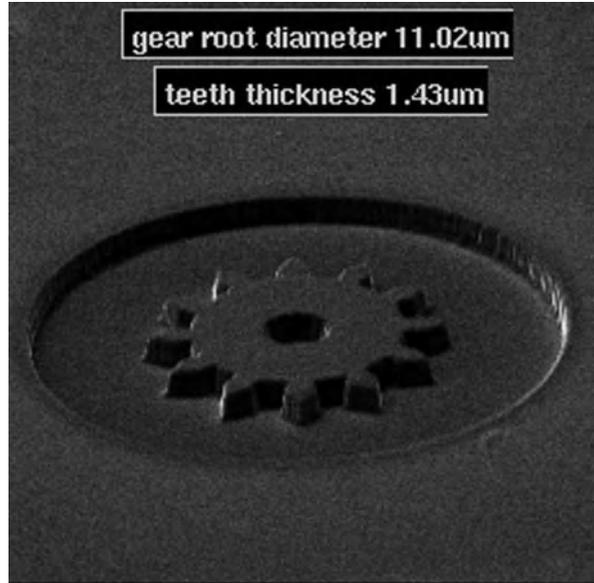
edge. These shapes are difficult to create directly using the FIB workstation's bitmap
editor. It is worth pointing out that the sidewall of the 2D patterns is not ideal
because of the large depth milled and the high fluence needed, for this leads to
a sloping sidewall because of both the redeposition effect and the Gaussian beam
profile. The greater the aspect ratio of the designed structures, the greater the slant
angle of the machined 2D patterns will be. Another example is seen in Fig. 6, which
shows a dragon pattern embedded on a human hair. The pattern was produced from
a bitmap pattern in binary *.XBM format [13].

2.2 Three-Dimensional Machining FIB

An energetic ion can interact with a target surface in various ways. Depending on
the ion energy, the interaction can result in swelling, deposition, sputtering, redeposition, implantation, backscattering, or nuclear reaction (see chapter "Basis of Ion
Scattering in Nanoscale Materials"). However, some of the interactions are not completely separable and may lead to unwanted side effects that need to be understood
and avoided for a specific application. For milling applications, it is desirable that
the incoming ions interact only with the atoms at the surface or near the surface layer
of the target substrate and give rise to a collision cascade on the atoms. If the ion
energy (or momentum) is sufficient, the collision can transfer sufficient energy to the
surface atoms to overcome their surface binding energy (3.8 eV for Au and 4.7 for
Si), and the atom is ejected as a result (Box 2). This interaction is called sputtering
and is the governing effect in FIB milling. Because the interaction depends solely
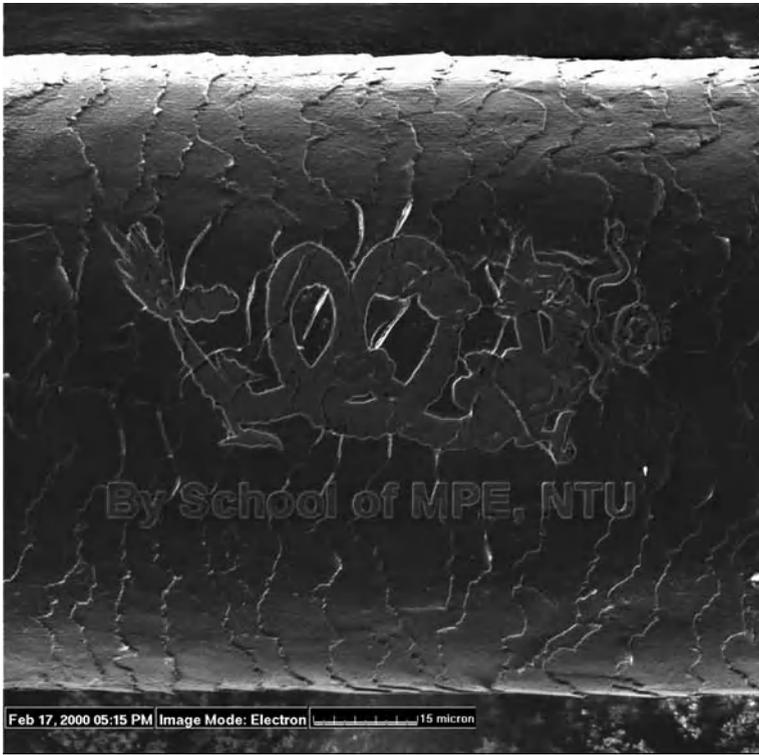
Fig. 6 A dragon pattern milled on a human hair with diameter of $60\,\mu m$ using the bitmap pattern method

on momentum transfer to remove the atoms, sputtering is a purely physical process. The sputtering yield, defined as the number of atoms ejected per incident ion, is a measure of the efficiency of material removal. The yield is normally in the range of 1–50 atoms per ion and depends on parameters such as the masses of the ions and target atoms, ion energy, direction of incidence to the surface of the target, target temperature, and ion flux. Initially, the sputtering yield increases as the ion energy increases, but the yield starts to decrease as the energy is increased, so the ions can penetrate deep into the substrate. At this stage of interactions, implantation or doping can take place in which the ions become trapped in the substrate as their energy is expended. As a result, the optimum energy for sputtering is between 10 and 100 KeV for most of the ion species used for milling. During sputtering, a portion of the ejected atoms or molecules are redeposited into the sputtered region and this redeposition makes it difficult to control the amount of material removed. In fact, the essence of FIB milling is to carefully control both the material sputtering and the redeposition, so that a precise amount of material can be removed. It should also be noted that since implantation by FIB mainly alters material properties and not material removal, only sputtering and redeposition will be examined in this chapter. The

removal of sample material is achieved using a high ion current beam. By scanning the beam over the substrate, an arbitrary shape can be etched. However, the above discussed sputter yield cannot be used directly to calculate the etch rate because, depending on the scanning style, redeposition occurs, which drastically reduces the effective etch rate. Furthermore, the sputtering yield is dependent on the angle of incidence, increasing roughly as $1/\cos(\theta)$, with θ the angle between the surface normal and the ion beam direction.

Two-dimensional machining by FIB is introduced in a number of papers and books. Here, we highlight two issues in this section: FIB programmable controlled machining, and gray-scale bitmap pattern controlled machining [1, 14–17].

2.2.1 FIB Programmable Controlled Machining

Fabrication of three-dimensional (3D) microstructures by use of the FIB technology is a potential application for a FIB in addition to its applications in modification and failure analysis of semiconductor devices. Vasile et al. introduced a depth control (also the pixel dwell-time changing) method for the fabrication of 3D microstructures using direct FIB milling [18–20]. In their method, they form the desired 3D structures by changing the sputter erosion via the dwell time at each pixel point. This allows a milling depth for different pixels. For each point, the required sputter yield (a function of ion beam incident angle) and dwell time were computed using a mathematical model [21]. The whole process is controlled by the FIB computer. Therefore, it needs a long time and many computer resources for the milling process.

Normal FIB milling can only realize two-dimensional (2D) structures (only the depth changes with milling time in a defined milling area, and lateral dimensions are fixed during milling). For the fabrication of 3D microstructure, the milling depth is different at different locations of a defined area. During the FIB milling, the stage is usually held stationary, and the ion beam scanned within the defined area by means of electrostatic beam deflection. Therefore, the whole process must be controlled point-by-point and line-by-line via the FIB equipped computer. There are two methods for the FIB direct writing. Here we present a new method for the 3D fabrication using FIB direct writing that employs a two-dimensional layer-by-layer approach [21].

Our method is to transfer the desired 3D cavities to many discrete 2D flat slices. In other words, we use many sequential 2D slices with small thickness to represent the desired 3D cavities that are to be milled. Then the milling of the 3D structures is achieved by milling many sequential 2D flat slices instead of the subparabolic cavities. For the total slices number N, a total number of milling steps N is needed. In our method, the sputter yield and dwell time are constant for milling the flat slices instead of the previously reported gradual geometric shape formation based on changing sputter erosion according to the dwell time at each pixel. For the formation of a given parabolic cavity by the FIB milling, we transfer the structure to discrete 2D slices with N set according to the maximum depth of the parabolic cavity; each slice has the same thickness, as shown in Fig. 7(a)–(d). The smooth sidewall profile transition permits the milled pixel addresses of the next slice to differ. In that way
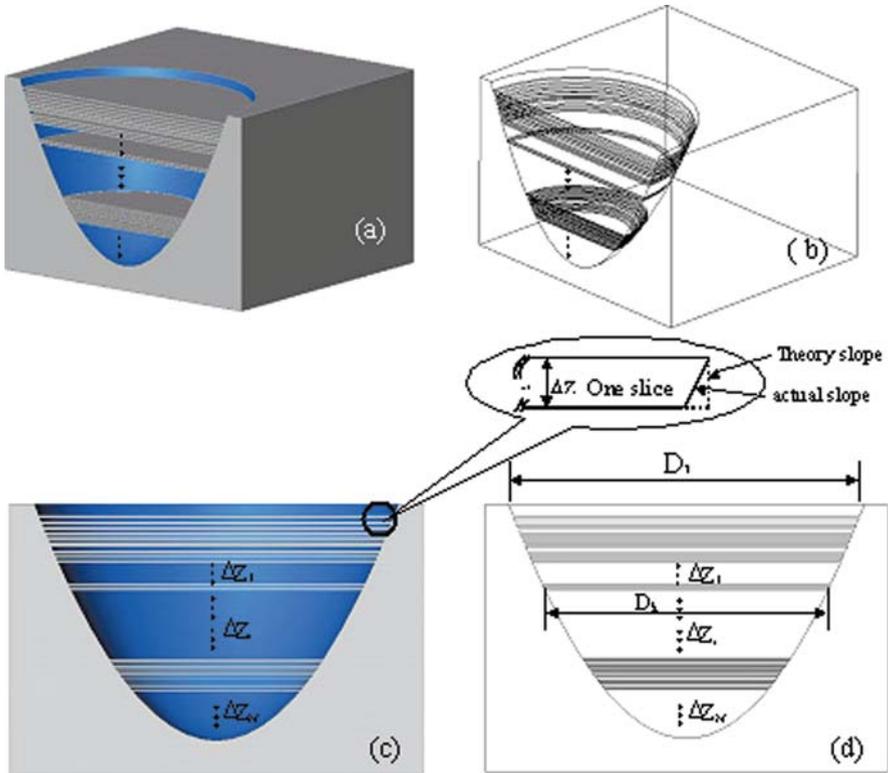
Fig. 7 Schematic cross-section diagram of (**a**), (**b**) 3D model, and (**c**), (**d**) 2D model shows that the milling process is divided into N slices or layers with different milling depth ΔZ_i for the profiles with different curvatures. In other words, we substitute the continuous curve with discrete data. Sputter yield and dwell time are maintained constant for each pixel in the defined area during milling of each slice

the bottom of the milled crater (the boundary between the two slices) is nearly planar. Assuming that the mean sputtering yield is a constant, it can be calculated for the case of a vertical ion bombardment. Because the 2D slices are flat, the sputter yield and dwell time are constant on the surfaces except for pixels at the periphery of the slices, for which the sputter yield differs because of the change in normal direction at the periphery, which causes varying of beam incident angles at the sidewall of the parabolic cavity. It will be shown by the following experimental results that the difference can be neglected due to its small influence on the contribution of the parabolic cavity formation. The number of slices, N, strongly depends on the maximum depth of the parabolic cavity. The steeper the sidewall, the more the slices number N will be. From the material parameters and the ion beam parameters, the dwell time at each pixel is solved to produce the milling depth, ΔZ The depth increment needed at the *(i, j)* pixel can be expressed as

$$\Delta Z_{ij} = \frac{St}{\eta} \iint \phi(x, y) f_{x,y}(x_i, y_j) dx dy \tag{9}$$

In Eq. (9), $\phi(x, y) f_{x,y}(x_i, y_i)$ gives the ion flux contribution to an arbitrary point (x_i, y_j) from a focused beam. The double integral gives the ion flux contribution at the point (x_i, y_j) from all points on the slice surface. S is the sputter yield for the ions in the normal incidence angle. It can be derived by the commonly used professional software, TRIM [22].

Equation (9) is still inconvenient for practical programming use because of functional limitations of some FIB machines, e.g., the Micrion 9500EX dual beam system does not support algorithm programs in the Unix operating system. Considering this, another model for the programming has to be adopted.

In practice, beam overlap is necessary for the FIB scanning line-by-line in a defined area. A negative overlap means that there is a space between adjacent pixel positions. A positive overlap means that the beam removes material from adjacent pixels. For a beam with diameter d_b and a step size of s, the overlap OL is related to s and d_b as $s = d_b (1-OL)$. The volume of the removed material, V, can then be written as [15]

$$V = \frac{SAN_{Ga}}{\rho N_A (1 - OL)^2} \tag{10}$$

where N_{Ga} is the number of incident ions per pixel (this can be transferred from ion fluence in practical cases), ρ the volume density of substrate material, N_A the number of atoms per pixel, and A the atomic mass.

The volume V can be transferred to the milling depth, ΔZ for parabolic cavity formation, as shown in Fig. 3(a)–(d). Each slice is a thin cylindrical layer with thickness ΔZ and diameter D_k $(k = 1 \dots N)$. Equation (9) can be rewritten as Eq. (11)

$$\Delta Z = \frac{SAN_{Ga}}{\rho N_A (1 - OL)^2} \frac{4}{\pi D_k^2} \tag{11}$$

The first step calculates the required N_{Ga} for the milling depth ΔZ of each slice offline, and arranges the milling sequences for a desired 3D cavity, then transfers the 3D cavity to N discrete 2D slices subsequently. The user inputs the geometric parameters and process parameters into a computer program (called MDDL in the Micron 9500EX FIB machine). A bitmap file with a circle pattern (for the fabrication of a parabolic cavity pattern) with the format of *.XBM and different diameters for the different milling steps is called by the MDDL program. Detailed information can be found in [13, 21]. The whole process for the milling of the desired 3D cavity can then be performed automatically by the program, which involves all parameters for the milling of the corresponding transferred 2D N slices.

The layer-by-layer method simplifies the previous mathematical model and controlling program. The layer-by-layer approach is simple, and the milling speed is

fast, because the tedious calculation of sputter erosion and dwell time were avoided in this new method.

2.2.2 Gray-Scale Bitmap Pattern Controlled Machining

Micro-optical elements (MOEs) are key components for optical communications, optical sensors, and biochemistry analysis systems. With integration and miniaturization, optical systems become smaller and smaller, such as micro-optical electronic mechanical systems (MOEMS). Currently, there are three main direct writing methods for direct fabrication of the micro-optical elements: laser direct writing (LDW) [23, 24], electron beam lithography (EBL) [25], and focused ion beam direct milling (FIBM). Proton beam writing (PBW) offers some superior characteristics and is rapidly emerging (see chapter "Proton Beam Writing: A New 3D Nano Lithographic Technique"). The FIB direct writing requires a programming function. Considering this, we reported a method of quasi-direct writing in [26]. This still has limitations for fabrication of circular symmetrical diffractive structures, e.g., Fresnel-like diffractive structures, because the ion beam cannot scan with variable spacing along radial directions. The ion beam can scan in circular track only for current commercial FIB machines. Here we introduce a new fabrication method. We call it the virtual gray-scale mask method because the "mask" is not a real mask. It is only a computer generated bitmap pattern with 256 grays in *.BMP format. The three-dimensional (3D) structures can be realized by means of calling the bitmap patterns with a file format of *.BMP in a defined area, and scanning the bitmap patterns by the FIBM.

The bitmap pattern can be designed by commercial software, e.g., Leica Qwin 96. Alternatively, users can write a computer program using $C++$, Visual Basic, or using MATLAB for generating a bitmap file with different gray scales. The corresponding converted gray scale is $G_i = 255 \times D_i/D_0$, where D_i is the relief height or depth at any location, and D_0 is the designed maximum relief height, as shown in Fig. 7, $G_i = 0$ for black color, and 255 for white color. As fabrication examples, we designed two bitmap patterns using the MATLAB programming. One is a blazed grating, and the other a diffractive lens. Figure 8 shows the two-dimensional (2D) profiles of the diffractive lens. After converting the designed 2D profiles into the corresponding gray-scaled profile (different radii correspond to different gray scales), the virtual gray-scale mask pattern can be plotted, as shown in Fig. 9. Figure 10 is the converted virtual gray-scale mask of the blaze grating as for the diffractive lens [27].

The blaze grating was fabricated using focused ion beam direct milling with the virtual gray-scale mask. The initial process parameters before the fabrication were 3 nA beam current, 30 keV ion energy, and 38 min milling time. The ion beam scan covered an area of $63.6 \times 63.6\,\mu m^2$. Other parameters were pixel size, overlap, beam current, and beam spot size of 15 nm, 50%, 100 pA, and 25 nm, respectively. Before starting the milling, the milling time was set according to calibration data (milling depth vs. milling time depends on different beam currents and substrate materials) and the maximum relief height (designed relief height D_0). The gray level, G_i, is
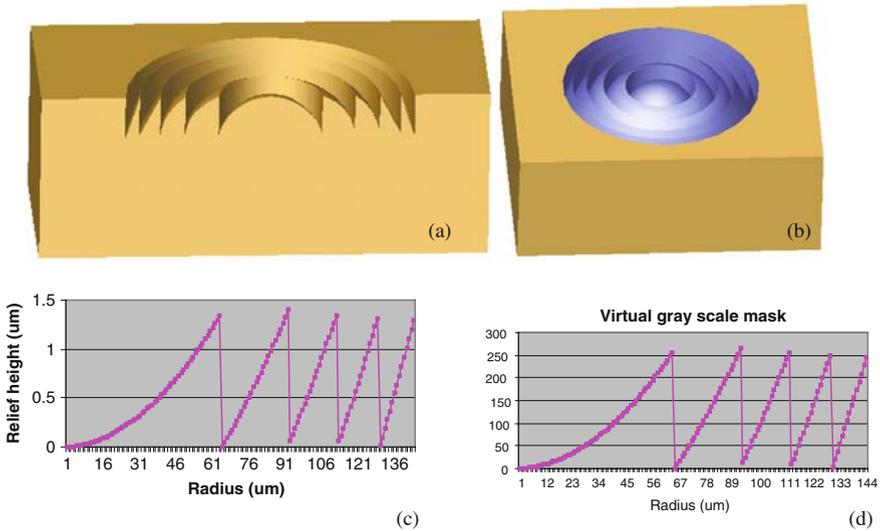
Fig. 8 Schematic of a designed diffractive lens with three-dimensional (3D) view. (**a**) Cross-section view; (**b**) overall 3D view; (**c**) designed 2D profile for the diffractive lens; and (**d**) transferred 2D profile of the diffractive lens with four annuli in terms of the corresponding gray scales ranging from 0 to 255 for further processing of MATLAB 5.3 software. Data sheet of gray scale as a function of radius will be used in the MATLAB 5.3 programming to plot the 2D gray-scale pattern of the lens
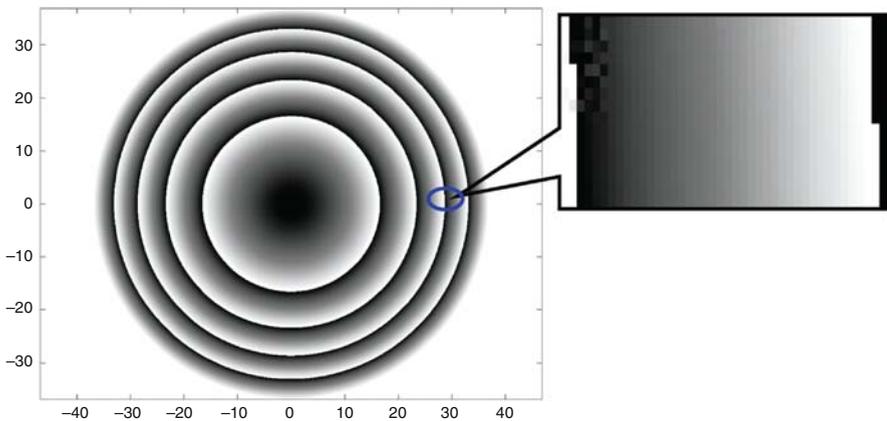
Fig. 9 Converted virtual gray-scale mask for the designed diffractive lens by use of MATLAB 5.3 programming. Function of gray scale vs. radius, $G_i(r)$, is used to calculate the gray scale corresponding to each point of the 2D mask pattern in the recycle function in the programming. Then the pattern is plotted with the plotting command in the MATLAB 5.3

Fig. 10 Converted virtual gray-scale mask of the blaze grating using the same method as that illustrated in Fig. 3, through MATLAB 5.3 for FIB machine use with period and blaze angle of $2.0\,\mu$m and $10°$, respectively
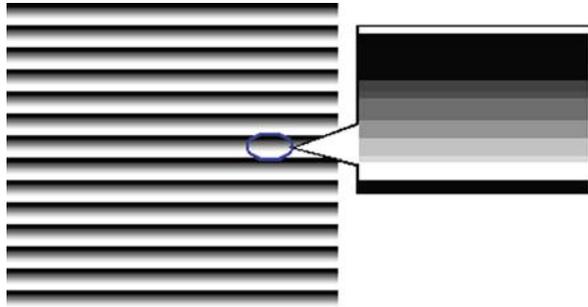
a relative value to the milling depth. For the diffractive lens milling, the setting time is 49 min for the maximum relief height $D_0 = 1.34\,\mu$m. The discrete step is $0.125\,\mu$m during the transformation, which is small enough for FIB writing. The total number of 2D patterns N was 2883. For the FIB writing, the operating software converts the virtual gray-scale mask to a corresponding milling depth proportional to the maximum designed depth. The conversion accuracy depends on the beam current and beam spot size used. In principle, it is similar to the reported method of gray-scale mask. But actually, it is only a bitmap pattern instead of a real mask that provides flexibility of the FIB direct writing.

The milling experiments were carried out by a new FIB machine (Quanta 200 3D Dual Beam System from FEI Company) with a liquid gallium ion source, integrated with a scanning electron microscope (SEM). This machine uses a focused ion beam with energy of 5–30 keV, a probe current of 1 pA–20 nA, and beam limiting aperture size of 15–350 μm. For the smallest beam currents, the beam can be focused down to as small as 10 nm in diameter at full width and half maximum (FWHM). The stage can be translated in the *X, Y,* and *Z* directions, and rotated in $360°$. It can also be tilted from $-10°$ to $60°$. During writing, the stage is stationary, and the ion beam is scanned on the sample surface in raster or serpentine mode, depending on the application. The ion beam is focused on the sample surface at normal incident angle.

We designed a micro-diffractive lens with continuous relief for focusing with a *NA* value and annulus number of 0.16 and 4, respectively. Such a lens can be used for integration with multi-mode fibers. For the micrograting, we selected $+1$ as the main diffraction order. Angular dispersion in $+1$ order was 0.37 at a wavelength of 1550 nm. In addition, a blaze grating with period and blaze angle of $2.0\,\mu$m and $10°$, respectively, was designed for the fabrication using our virtual gray-scale mask.

Figure 11 shows a FIB image of the blazed grating fabricated using the FIB with the virtual gray-scale mask in a defined area of $15 \times 15\,\mu\text{m}^2$ on a silicon wafer. It can be seen that the surface transition is smooth. This was characterized by the use of an atomic force microscope (AFM from DI Company, NanoScope IIIa) in tapping mode, as shown in Fig. 12(a) and (b). The scanner was calibrated against a 160 nm height standard. The measured period and blazed angle of the grating is $1.98\,\mu$m and $9.17°$, respectively, which meet our design parameters. Figure 13 is a

x 2.000 μm/div
z 800.000 μm/div

Fig. 12 AFM measurement results of the fabricated blaze grating using focused ion beam direct milling with virtual gray-scale mask. (**a**) 3D image, and (**b**) 2D profile

SEM micrograph of the diffractive lens fabricated using the FIB direct milling in terms of the virtual gray-scale mask on a silicon wafer.

The attainable writing resolution depends on the ion beam spot size which ranges from 10 to 200 nm. Selection of the beam spot size depends on the dimensions of the MOEs. For our diffractive lens with an outer diameter of $63.6\,\mu m$, we used a beam current of 3 nA, which corresponds to a beam spot size of 80 nm. The pixel overlap during the beam scanning is 50% in both X and Y.

2.2.3 Self-Organized Formation Under Ion Beam Bombardment

For some materials, ripples can be spontaneously generated by self-organization during ion beam bombardment. The feature sizes of the structure are governed by ion energy, beam incident angle, and scan time. Once the parameters are determined,
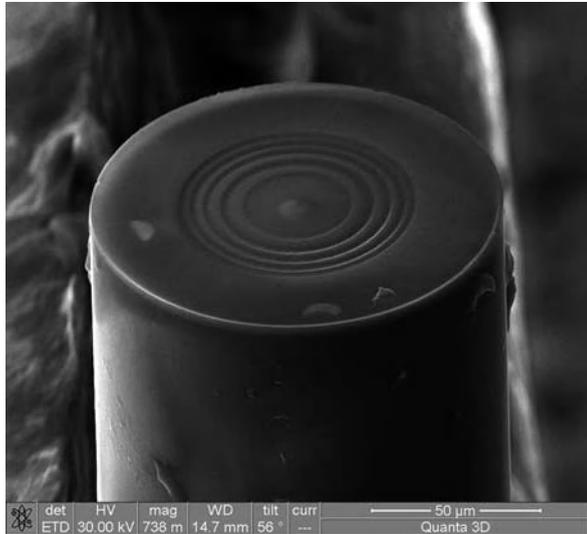
they are fixed for the whole scanning process. The ripples with nanoscale can be used as quantum wires in semiconductor lasers, and nanowires for biomedical sensing. A typical theoretical model describing the rippling is the BH model [28]. The wavelength and amplitude of the ripples strongly depend on ion beam energy, scan speed, ion beam incident angle, and ion beam projection direction. The ripple orientation depends on the angle of incidence. For a long bombardment, a nonlinear model has been developed for describing the evolution of surface morphologies such as roughening, coarsening, and saturation of ripples by destroying the ripples and generating a new rotated ripple structure [29].

As an example, it has been reported that ripples can be generated on Si (001) surface after bombardment with a focused 30 keV Ga^+ beam at 52° incident angle at room temperature (Fig. 14(a)). The ripple wavelength is perpendicular to the projected ion beam direction, and is independent of raster direction of the beam and the overlapping percentage between adjacent FIB points. The orientation and
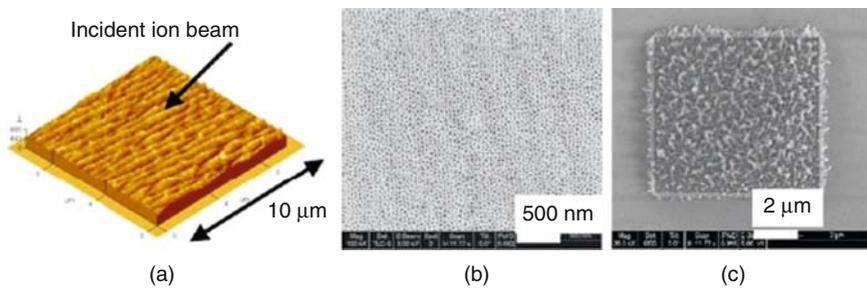
Fig. 14 FIB-induced surface features: (**a**) Ripple formation on Si (001), (**b**) nanodot formation on Ge (001), and (**c**) fiber formation on Ge (001)

Fig. 15 Highly ordered Ga droplet formation on the surface of GaAs induced by FIB under 5 keV at incident angle 35°. Inset showing the Fast Fourier Transformation (FFT) [31]
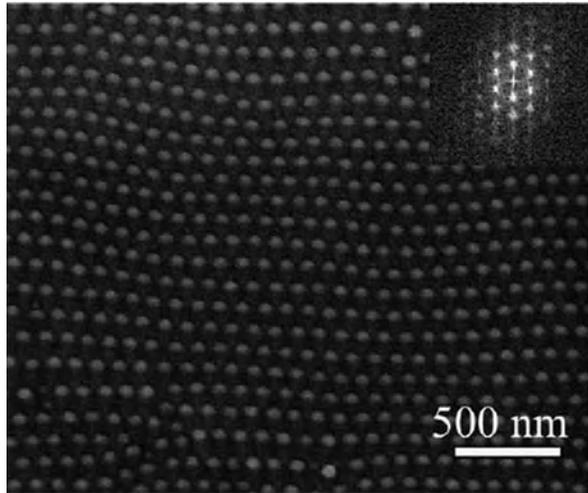
wavelength of ripples, as proposed by the BH model, is due to a balance between curvature dependent roughening and surface smoothing. The ripple appears to be more regular under FIB bombardment for an unfocused ion beam under the same conditions on Si surface [30]. However, on the surface of Ge, nanodots have been found for low irradiations (3.75×10^{15} ion cm^{-2}) for a focused 30 keV Ga$^+$ beam with normal bombardment (Fig. 14(b)). At high enough fluences, "nanospiked" surfaces with embedded fibers are produced (Fig. 14(c)). Recently, highly ordered Ga nanodroplet formation on the GaAs surface induced by low energy FIB at off-normal bombardments has been reported (Fig. 15) [30]. Due to implanted Ga atoms and preferential sputtering of As atoms, a larger amount of Ga is generated on the surface of GaAs during bombardment. These Ga atoms have a high mobility and can aggregate into Ga droplets. However, undesired sputtering can occur on the surface of both droplet and substrate, and there is a limitation on the growth of droplets. Under off-normal bombardment, a balance between anisotropic loss of atoms from the surface of droplets due to sputtering and anisotropic supply of atoms on the substrate surface due to preferential sputtering of As gives rise to the well-organized patterns [31].

2.3 Ga Ion Implantation and Radiation Damage

Most semiconductor devices are now doped using a broad area or a flood beam ion implantation of patterns which are delineated by various masking materials such as photoresist, metal films, and SiO$_2$ or Si$_3$N$_4$. The implantation process is highly automated and production systems are capable of handling more than 200 wafers per hour. Implantation energies range from < 30 keV in predeposition implanters, through 150–200 keV for medium and high current implanters, to several MeV for

the latest high energy systems. Focused ion beam implantation (FIBI) can realize localized ion implantation on the surface of devices or chips. The implanted area can form a protective layer for chemical etching. This is useful for MEMS device and diode laser chip fabrication [32, 33]. However, for normal commercial FIB machines, the ion energy is restricted to a maximum of 50 keV, which limits its use for localized implantation to shallow implants only. Numerical simulations revealed that the implantation depth is ~20 nm for typical Ga^+ ion bombardments with energy of 30 keV, as shown in Fig. 16 [34].

Figure 16 shows the calculated projected range (mean implanted depth) of Ga^+ vs. different ion energies estimated using TRIM 2000 [23]. It can be seen that the implantation depth for the fused silica surface is 20.8, 25.7, and 31.3 nm for the ion energy of 30, 40, and 50 keV, respectively. As an example, we carried out FIB milling experiments on a fused silica surface with ion energies of 30, 40, and 50 keV, respectively. The average concentration of the implanted Ga^+ measured by the electron dispersion X-ray spectrometer (EDX) is 9.8%, 12.3%, and 15% for the ion energy of 30, 40, and 50 keV, respectively, as shown in Fig. 17. Penetration depth of the X-ray is around 30 nm under the sample surface, that just matches with the implantation depth of the Ga^+ at 50 keV.

Figure 18 shows the relationship of refractive index versus wavelength before and after the FIB milling process measured with an ellipsometer (Mizojiri Kogaku) equipped with He-Ne laser. For the bulk material of fused silica, the measurement of extinction coefficient k is strongly affected by the presence of water or OH absorption in the samples. It makes the determination of the intrinsic k values extremely difficult for the wavelength in the range $3.7 \geq \lambda \geq 0.21$ [35]. Thus no k values were given in this figure. It can be seen that the refractive index increased around 0.5 after the FIB implantation with Ga^+ ions. The higher the ion energy used, the greater the
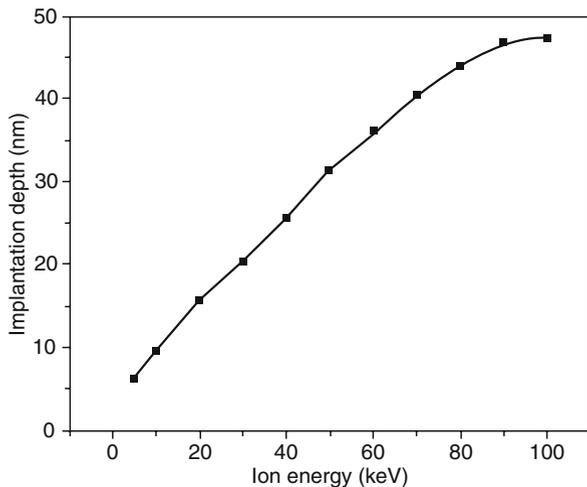
Fig. 16 Simulation results of ion implantation depth vs. ion energy in longitudinal direction. The relationship between implantation depth and ion energy is linear in the range from 5 to 80 keV
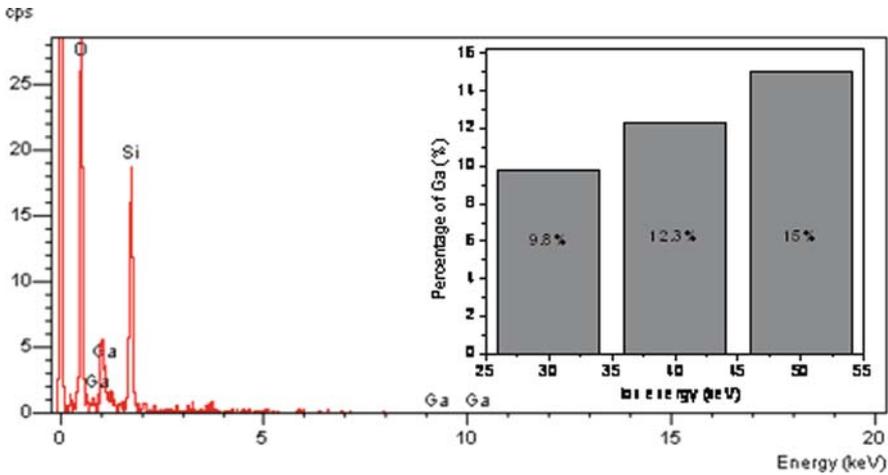
Fig. 17 Results of chemical elements analyzed by EDX. The percentage of Ga implanted in the quartzes were 9.8%, 12.3%, and 15% for the ion energy of 30, 40, and 50 keV, respectively

resulting refractive index. Clearly, increasing the energy of the Ga$^+$ implantation increases the refractive index.

Samson et al. [36] put forth a theoretical model for the change of permittivity ε after Ga ion implantation. The dielectric constant of Ga$^+$-doped fused silica, according to the Drude model $\varepsilon_{\text{doped}}$ can be approximated by the following expression:
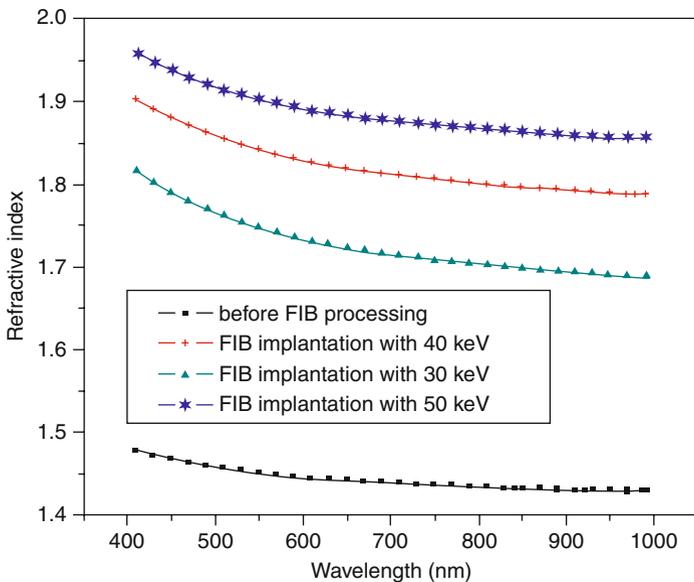
Fig. 18 Refractive index of the fused silica vs. wavelength before and after ion beam bombardment with different ion energy ranging from 30 to 50 KeV

$$\varepsilon_{doped} = \varepsilon_{pure}\left(1 - \frac{\omega_p^2}{\omega\left(\omega + i\frac{1}{\tau_{imp}}\right)}\right) \tag{12}$$

$$\varepsilon_{pure} = \varepsilon_{Si} + \left(\frac{\alpha}{\lambda}\right)^2 + \frac{\gamma}{(\lambda/\beta)^2 - 1} \tag{13}$$

$$\omega_p = \sqrt{\frac{e^2 N}{\varepsilon_r \varepsilon_0 m_{eff}^c}} \tag{14}$$

$$\tau_{imp} = \frac{16\pi\sqrt{2m_{eff}^*}}{NC_{BH}}\left(\frac{\varepsilon_r \varepsilon_0}{Ze^2}\right)^2 E^{3/2} \tag{15}$$

where τ_{imp} is the collision time, N the doping concentration of the sample, e the electron charge, ε_r the relative permittivity, ε_0 the vacuum permittivity, m_{eff}^* the effective optical mass of the charged carrier, m_{eff}^C the electron mass, Z the charge state of the impurities, C_{BH} the Brooks-Herring correction term, and E the ion energy.

3 FIB Direct Deposition

3.1 Mechanisms of Material Deposition

FIB can be used for the localized maskless deposition of both metal and insulator materials. The principle is based on chemical vapour deposition (CVD) and the reaction process is similar to, for example, laser-induced CVD [37, 38]. The main difference from laser-induced CVD is the better resolution and lower deposition rate with FIB. The metals that can be deposited on commercially available machines are mainly platinum (Pt) and tungsten (W). In the case of W, the organometallic precursor gas is $W(CO)_6$. An insulating material, SiO_2, can be deposited with 1, 3, 5, 7-tetramethylcyclotetrasiloxane (TMCTS) and oxygen (O_2), or alternatively water vapor (H_2O), as precursors. The precursor gases are sprayed on the surface by a fine needle (nozzle), where they adsorb. In a second step, the ion beam decomposes the adsorbed precursor gases and the volatile reaction products are desorbed from the surface and pumped away by the vacuum system, while the desired reaction products (Pt, W, or SiO_2) remain fixed on the surface as a thin film. Generally the deposited material inevitably contains organic contaminants as well as Ga ions and atoms from the ion beam. Insulator deposition also has the additional complication of possible charging problems and metal contamination by the primary beam (Ga ions are implanted into the deposition substrate). Normally, the overlap between the neighboring beam spots is 5–10 times larger for deposition compared to milling because of so-called overspray (deposition outside the irradiated area).

3.2 Three-Dimensional Formation by Deposition

A microcylindrical lens is used here as an example to demonstrate the mechanism of the FIB direct deposition. These lenses are useful for collimating or focusing light beams for laser-beam shaping, beam interconnects, and micromedical device applications. Conventional fabrication methods are e-beam and g-line optical lithography [34, 35]. The FIB has been used for micromachining since the early 1990s [39–45] by Kung, Shank, and coworkers, who used it to mill multilevel diffractive optical elements [46]. FIB has great potential for micromachining optical elements, especially because of its deposition capability. The patterned deposition of SiO_2 makes it possible to use it for post-processing of surface-relief microstructures or to add specific microoptical elements at accurately positioned locations of a microdevice, such as a micro-electromechanical system (MEMS) or micro-optoelectromechanical system (MOEMS).

By changing the dwell time under different beam-limiting aperture sizes and controlling the actual beam current error to be less than 10% of the corresponding theoretical value, different deposition results for SiO_2 were obtained with precursor gases of 1,3,5,7 tetramethylcyclotetrasiloxane (TMCTS) and H_2O for the SiO_2 microlens deposition. The accelerating voltage used was 50 kV, and the chamber base pressure was maintained at 9.1×10^{-1} Pa during the whole deposition process. The continuous profile of the microcylindrical lens was divided into many discrete thin layers to form the 3-dimensional profile of the lens. Each layer is rectangular, with a different size for each different layer. The microcylindrical lens is then composed of many superposed deposited rectangular SiO_2 films with different sizes corresponding to the curvature of the lens. Figure 19 is an SEM image of the microcylindrical lens obtained by SiO_2 deposition.

Two key factors determine the deposition quality: the implanted Ga^+ content and the stoichiometry of Si and O. The former is problematic in both the milling and deposition process because of unavoidable ion implantation. These factors affect the optical transmission of the microlens transmission directly. For deposition of the SiO_2 film, the beam is stopped to some degree during the process of decomposition and adsorption, leading to less Ga^+ to be implanted than is the case for milling.

The H_2O and TMCTS pressures are key parameters for the O_2 to Si stichometry during the deposition of the SiO_2 film, which determines the optical transmission of the SiO_2 material. The transmission will be ideal at a stichometric SiO_2 composition (Si: 46.74 w.%, O: 53.26 w.%), which can be obtained by adjusting the H_2O and TMCTS pressures. If the chamber base pressure and temperature are too high, contamination leads to an irregular pattern. If the pressure is lower than 10^2 Pa and the temperature is lower than 40 °C, the precursor gas feed is insufficient for decomposing. Sputtering also takes place during the deposition process, and milling of the substrate's surface will also occur if the precursor pressure is low. Typical values are 1.5–5 $\times 10^2$ Pa and 50 °C–52 °C for the precursor partial pressure and temperature, respectively.

As an application example, a microcylindrical lens designed for collimation is directly deposited on the emitting facet of a laser diode, as shown in
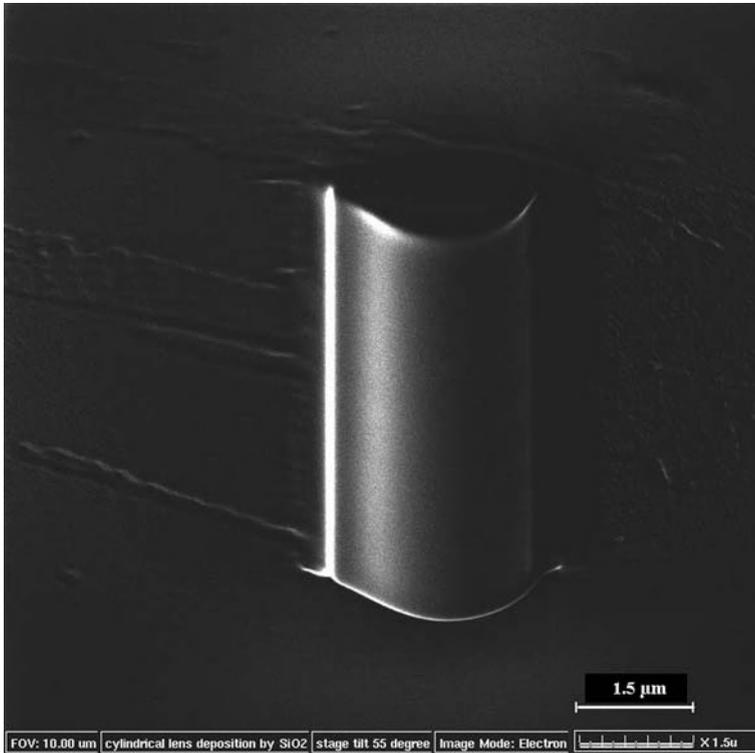
Fig. 19 SEM microphotograph of microcylindrical lens with numerical aperture (NA) of 0.4 by SiO$_2$ deposition imaged at 30° view angle

Figs. 20 and 21. The controlled SiO$_2$ deposition process was realized by programming of the FIB machine (Micrion 9500EX). Using a single-mode fiber with core diameter of 10 μm as a test prototype, the coupling efficiency of the compact and miniaturized system could be as high as 80%. The measured spot size is about
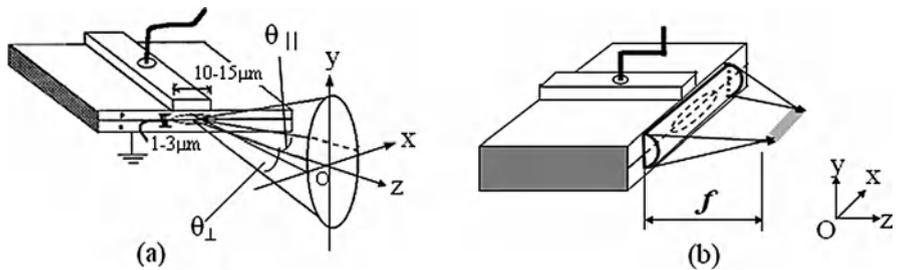
Fig. 20 Schematic of laser diode structure. (**a**) Diagram of original laser diode; (**b**) integrated microcylindrical lens with the laser diode. The microlens with size $50 \times 5 \times 0.85 \, \mu m^3$ covered on the emitting facet
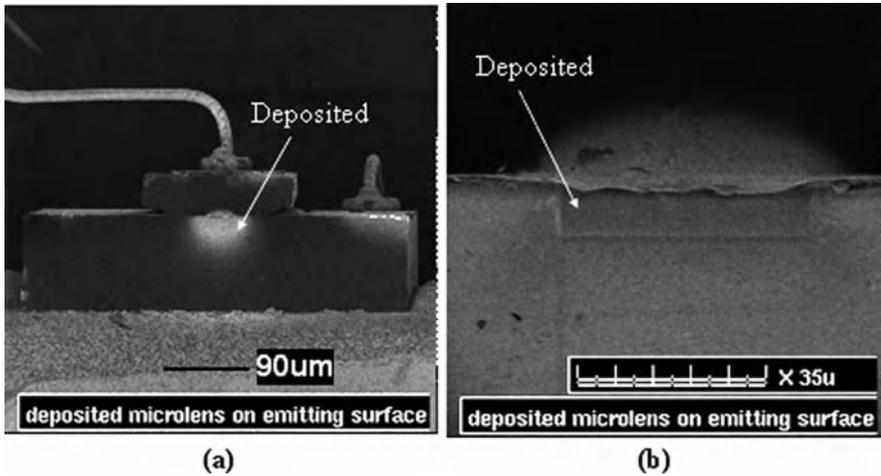
Fig. 21 Micrograph of deposited microcylindrical lens on emitting surface of laser diode. (**a**) 3D view of integrated laser diode (**b**) top view of microlens

5.4 μm and 9.0 μm (at $1/e^2$: 13.5%) in the meridional and sagittal planes respectively. The coupling efficiency measured here could be enhanced by the use of a suitable antireflection coating and decreasing microfabrication error and measurement misalignment error. The measured far field angle is 2.1° and 31° with and without the lens, respectively. For more information, see [47].

4 Future Perspective

As a key nanofabrication tool, FIB will still play a dominant role for nanotechnology and nanosciences in the future, especially in the field of prototyping of nanophotonic devices, which has been an emerging research area in recent years [48–52]. However, scientists are focusing on improving and developing new functional FIB machines because of the shortages of small fabrication areas and slow machining speeds for the current commercial FIB machines. One of the main points is to design a special ion column which can generate an array of focused ion beams. It means that a large fabrication area can be formed by use of the FIB array which mill each individual pattern simultaneously and therefore the large fabrication area is generated by spelling of the individual patterns side-by-side and line-by-line. Large area and high speed fabrication can be realized accordingly by this approach. This function is significant for devices and sensor fabrication in nanophotonics and biosensing. The other trend is to modify the FIB machine to be a "nanolathe." Out-of-plane symmetrical micro/nanostructures can be machined using this nanolathe. It will be popular in precision engineering, nano-electromechanical systems (NEMS), etc.

References

1. J. Melngailis, J. Vac. Sci. Technol. B 5, 469 (1987).
2. D. K. Stewart, A. F. Doyle and J. D. Casey, Proc. SPIE 2437, 276 (1995).
3. S. Reyntjens, D. De Bruyker and R. Puers, *Proceedings of 1998 Microsystem Symposium* (Delft, the Netherlands, 1998) p. 125.
4. B. W. Ward, N. P. Economou, D. C. Shaver, J. E. Ivory, M. L. Ward and L. A. Stern, Microcircuit modification using focused ion beams. Proc. SPIE 923, 92 (1988).
5. J. Glanville, Solid State Technol. 32, 270 (1989).
6. D. K. Stewart, L. A. Stern, G. Foss, G. Hughes and P. Govil, Proc. SPIE 1263, 21.
7. J. F. Walker, J. C. Reiner and C. Solenthaler, *Proceedings of Microscopy Semiconductor Materials Conference* (Oxford, 20–23 March 1995) p. 629.
8. G. J. Athas, K. E. Noll, R. Mello, R. Hill, D. Yansen, F. F. Wenners, J. P. Nadeau, T. Ngo and M. Siebers, Proc. SPIE 3223, 198 (1997).
9. J. H. Daniel and D. F. Moore, Sensors Actuat. A 73, 201 (1999).
10. J. Brugger, G. Beljakovic, M. Despont, N. F. de Rooij and P. Vettiger, Microelectron. Eng. 35, 401 (1997).
11. S. Reyntjens and R. Puers, J. Micromech. Microeng. 11, 287 (2001).
12. G. Benassayag, P. Sudraud and B. Jourffrey, Ultramicroscopy 16, 1 (1985).
13. Y. Fu, N. K. A. Bryan, Int. J. Adv. Manuf. Technol. 16(8), 600–602 (2000).
14. P. D. Prewett (Philip Doughty) edited, *Focused Ion Beams from Liquid Metal Ion Sources* (Research Studies Press, Taunton, UK, 1999).
15. J. Orloff edited, *Handbook of Charged Particle Optics* (CRC Press, Boca Raton, 1997).
16. S. Reyntjens and R. Puers, J. Micromech. Microeng. 11, 297–300 (2001).
17. A. A. Tseng, Micromech. Microeng. 18, R15–R34 (2004).
18. M. J. Vasile, J. Xie and R. Nassar, J. Vac. Sci. Technol. B 17, 3085 (1999).
19. R. Nassar, M. Vasile and W. Zhang, J. Vac. Sci. Technol. B 16, 109 (1998).
20. M. J. Vasile, R. Nassar and J. Xie, J. Vac. Sci. Technol. B 16, 2499 (1998).
21. Y. Fu and N. K. A. Bryan, J. Vac. Sci. Technol. B 22, 1672 (2004).
22. A commonly used software for simulation and analysis of ion beam processing. It can be found at: http://www.srim.org/srim/srimlegl.htm.
23. A. Y. Smuk and N. M. Lawandy, Opt. Lett. 22, 1030 (1997).
24. M. T. Gale, M. Rossi, J. Pedersen and H. Schutz, Opt. Eng. 33, 3556 (1994).
25. T. Fujita, H. Nishihara and J. Koyama, Opt. Lett. 6, 613 (1981).
26. Y. Fu, N. K. A. Bryan and W. Zhou, Opt. Exp. 12, 1803 (2004).
27. R. M. Bradley and J. M. E. Harper, J. Vac. Sci. Technol. A 6, 2390–2395 (1988).
28. R. Cuerno and A.-L. Barabasi, Phys. Rev. Lett. 74, 4746 (1995).
29. A. Cuenat and M. J. Aziz, Mat. Res. Soc. Symp. Proc. 696 (2002).
30. Q. M. Wei, J. Lian, W. Lu and L. M. Wang, Phys. Rev. Lett. 100, 076103 (2008).
31. J. Xu and A. J. Steckl, Appl. Phys. Lett. 65(16), 2081 (1994).
32. S. Eshlaghi, C. Meier, D. Suter, D. Reuter and A. D. Wieck, J. Appl. Phys. 86, 6605–6607 (1999).
33. Y. Fu and N. K. A. Bryan, Appl. Phys. B, 80(4), 581–585, 6 April 2005.
34. E. D. Palik edited, *Handbook of Optical Constants of Solids*, Academic Press, Inc., Boston (Harcourt Brace Jovanovich, Publishers), 1985.
35. J.-S. Samson, G. Wollny, E. Brundermann, A. Bergner, A. Hecker, G. Schwaab, A. D. Wieck and M. Havenith, Phys. Chem. Chem. Phys. 8, 753–758 (2006).
36. Y. Fu and N. K. A. Bryan, Opt. Eng., 44(12), 128002 (2005).
37. G. Thornell and S. Johansson, J. Micromech. Microeng. 8, 251 (1998).
38. S. Johansson, J.-A. Schweitz, H. Westberg and M. Boman, J. Appl. Phys. 72, 5956 (1992).
39. A. Stemmer, H. Zarschizky, E. Knapek and F. Mayerhofer, Microelectron. Eng. 27, 171 (1995).

40. A. Stemmer, H. Zarschizky, F. Mayerhofer, G. Lefrance, W. Gramann and P. Galloway, SPIE 1732, 77 (1992).
41. H. Zarschizky, A. Stemmer, E. Knapek and F. Mayerhofer, Microelectron. Eng. 27, 175 (1995).
42. I. M. Templeton and H. G. Champion, J. Vac. Sci. Technol. B 13, 2603 (1995).
43. A. Kalburge, A. Konkar, T. R. Ramachandran, P. Chen and A. Madhukar. J. Appl. Phys. 82, 859 (1997).
44. L. Bischoff, J. Teichert and E. Hesse, Microelectron. Eng. 27, 351 (1995).
45. J. F. Walker, D. F. Moore and J. T. Whitney, Microelectron. Eng. 30, 517 (1996).
46. R. M. Bradley and J. M. E. Harper, J. Vac. Sci. Technol. A 6, 2390–2395 (1988).
47. Y. Fu and N. K. A. Bryan, J. Vac. Sci. Technol. B 19(4), 1259–1263 (2001).
48. W. L. Barnes, A. Dereux and T. W. Ebbesen, Nature 424, 824–830 (2003).
49. L. Martin-Moreno, F. J. García-Vidal, H. J. Lezec, A. Degiron and T. W. Ebbesen, Phys. Rev. Lett. 90, 167401 (2003).
50. Y. Fu, W. Zhou, L. E. N. Lim and C. Du, Appl. Phys. B 86(1), 155–158 (2007).
51. Y. Fu, W. Zhou, L. E. N. Lennie, C. Du and X. Luo, Appl. Phys. Lett. 91(6), 061124 (6 August 2007).
52. Y. Fu, W. Zhou, L. E. N. Lim and C. Du, Opt. Eng. 45(10), 108001, 17 October 2006.

Box 8: Sample Preparation for Transmission Electron Microscopy Using a Focused Ion Beam

Lumin Wang and Yongqi Fu

The most common conventional preparation techniques for cross-sectional transmission electron microscope (TEM) samples are a combination of mechanical grinding, polishing, and ion milling. These techniques suffer from the fact that they are very time consuming and that it is extremely difficult to do site-specific cross-sections. Conventional ion-milling techniques for TEM specimen preparation are essentially site blind, even with the so-called precision ion polishing systems (PIPS). Thus, it is left to chance whether the specimen detail of interest is thin enough for TEM-imaging. Even if subsequent ion milling of TEM specimens is possible in extending the thin areas, it is a tedious and mostly uncontrolled procedure. Because continuous high-resolution control of the ion milling process is not possible, the specimen areas of interest may easily be destroyed.

The advantage of using a focused ion beam (FIB) for TEM sample preparation is that it is less time consuming, and more importantly has the possibility of being site-specific. The sample surface is imaged with the FIB before cutting to define the area of interest. Then the sample is milled and polished with predefined milling patterns. In the final step the sample can be cut out of the substrate and transferred to a TEM grid. Generally, as only the sample surface is imaged, even the standard FIB cross-sectioning is basically a blind process without the possibility of simultaneous high resolution monitoring of the milling process, as the area of interest is easily destroyed.

Figure 1 shows schematic illustrations and application examples for TEM sample preparation using FIB. The detailed description can be found in [1]. In TEM sample preparation, real time SEM imaging allows for very tight control of the sample thickness/transparency, and the danger of destroying the fine lamella sample is reduced to a minimum. This FIB lift-out technique allows thin membranes to be extracted from bulk material, which saves sample pre-thinning time and has

L. Wang (✉)

Department of Nuclear Engineering & Radiological Sciences, Department of Materials Science & Engineering, Electron Microbeam Analysis Laboratory (EMAL), University of Michigan, Ann Arbor, MI 48109-2104, USA

e-mail: lmwang@umich.edu

291
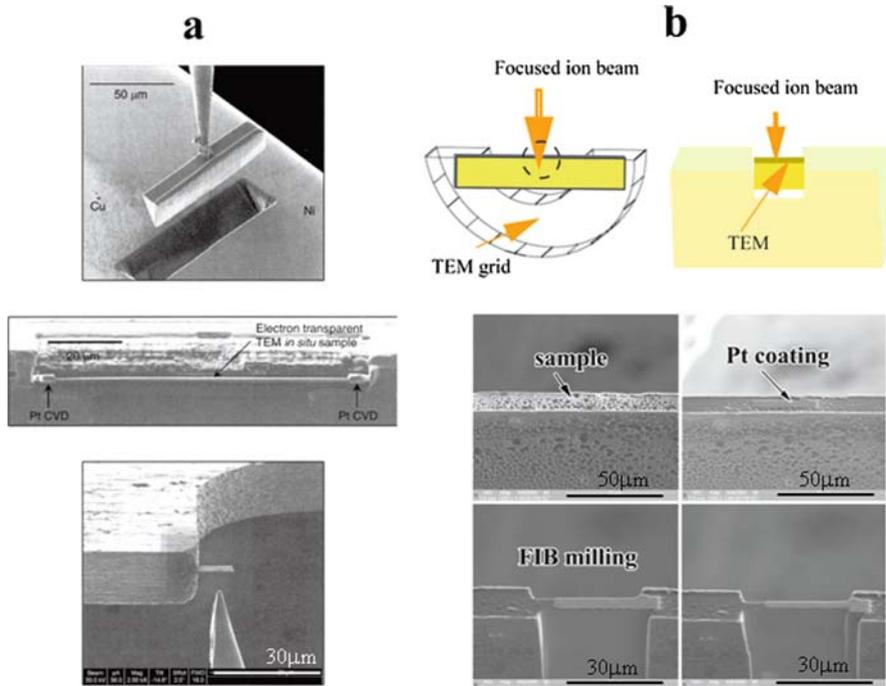
Fig. 1 Schematic illustration of process for TEM sample preparation and application examples: (**a**) example for liftout of a sample piece using a nanomanipulator inside the FIB chamber, [49]. *Top*: a piece of sample is mounted to the nanominipulator needle by Pt deposition, *middle*: mounted to a TEM grid, *bottom*: a shorter lamellae is mounted to the side of TEM grid. (**b**) Schematic illustration of H-bar technique (*top*) and example of thinning procedure (*bottom* four images)

proved very successful for the preparation of site-specific cross-sections and planar samples. To a large extent TEM sample preparation can be automated by using scripts and macros. However, the best compromise concerning time and accuracy is achieved if different samples are pre-thinned automatically overnight to a thickness of about 1 μm and then polished manually under high resolution SEM observation. For the final investigation of the TEM lamellae, energy filtered analysis enables optimum image contrast, even for thicker samples or heavy ones, because the chromatic smearing from inelastically scattered electrons can be eliminated. By selecting appropriate energy loss windows, element-specific contrast enhancement and quantitative determination of 2D composition distributions can be achieved. Moreover, the small probe size of advanced TEM instruments allows spectroscopic analysis in scanning TEM (STEM) or nanoprobe modes to reveal composition and interface properties on the subnanometer scale.

Reference

 1. J. Mayer, L.A. Giannuzzi, T. Kamino, and J. Michael, MRS Bull. 32, 400 (2008).

Box 9: Integrated Circuit Chip Modification Using Focused Ion Beams

Yongqi Fu

In this box, we will present the use of the micromachining and microdeposition capabilities of *Focused Ion Beams* (FIB) technology for modification of *Integrated Circuits* (IC) chips. The FIB repair of microcircuits is reviewed in detail by Sudraud et al. [1], who discussed several key applications. As the degree of circuit integration increases, the number of defects, such as unwanted connections and breaks, increases. This may be associated with recrystallization effects in the surface of the substrate, which was amorphized during the ion implantation process stage, as well as lithographic defects and aerosol particle deposition. FIB track-cutting by milling is already being seriously considered for commercial application in the repair of highly specialized, expensive custom ICs. Beyond the relatively simple operation of track cutting, a more sophisticated procedure for fine trimming of resistors presents an intriguing prospect for future highly precise IC technology. Just as unwanted tracks can be removed, absent connections can be corrected by FIB microdeposition. Commonly used deposition materials for the IC chip modification are Au, W, Pt, and SiO_2. The most efficient approach is to produce conductive layers by FIB decomposition of gas directed to the surface of the IC from a nozzle dispenser. This FIB technique has considerably better resolution than laser-stimulated chemical deposition.

The deposition process is governed by the following parameters: (1) the rate at which the precursor gas is absorbed onto the sample substrate; (2) the rate at which the precursor is decomposed by the incident ion beam; and (3) the rate at which the newly deposited film is sputtered away by the incident beam. For a defined X and Y pixel space, the first step determines the amount of precursor absorbed during the loop time t_l. The second step determines the amount of precursor converted to metal during the pixel dwell time t_d. If t_d is not sufficiently large for monolayer coverage to occur, or if t_l is too long, then sputtering of the metal occurs.

Y. Fu (✉)
School of Physical Electronics, University of Electronic Science and Technology of China, Chengdu 610054, P.R. China
e-mail: yqfu@uestc.edu.cn

The direct deposition process is that a gaseous chemical compound containing the deposition material is flowed onto the specimen while it is being rastered by ion beam. Much of the gas dissipates into the work chamber and is removed by the vacuum pumps, but a monolayer is absorbed by the surface of the specimen. As the ions pass through this monolayer, they decelerate rapidly, which results in an energy transfer that dissociates the gas molecules and fixes the deposition material on the surface of the specimen. Many repetitions of the raster are needed to build up successive layers of the material until the desired thickness is obtained. The raster scan technique for deposition involves many iterations of a raster that precisely administers a relatively low ion fluence. If the ion dose is too low, the material will not adhere; if it is too high, the ion milling effect will dominate. Considering this, serpentine scanning is typically used to minimize the blanking tail effect that may cause line-width broadening. The image is better for a smaller aperture size and a lower beam current. For a large ion beam current, the scanning speed is fast because of the large beam spot size and steps X and Y. As a result, the source gas has not been fully decomposed by ion impacts on the scan area during the total deposition time, and the deposition quality will be worse.

Deposition rate can be influenced by (a) the adsorption rate, (b) the decomposition rate, and (c) the physical sputtering rate of the deposited species. The decomposition of the precursor molecules can be caused by (a) direct ion impact, (b) sputtering ion or atom impact, (c) secondary electrons emitted upon ion impact, or (d) thermal spikes. The deposition yield is seen to increase with increasing pressure and decrease with decreasing temperature.

Experiments have been done on the relationship among deposited height, gas flux, and gas source reservoir temperature. The total ion fluence was $4\,\text{nC}/\mu\text{m}^2$ and was kept constant throughout the experiment. The scanned area and beam current were maintained constant at $2 \times 10\,\mu\text{m}^2$ and $152\,\text{pA}$, respectively [2]. The relationship between gas flux and deposited height was an increasing exponential. The precursor gas flux increased with increasing reservoir temperature. However, this was limited by heating of the sample chamber, which degraded the vacuum and introduced contamination.

In IC modification it is often important to precisely determine the endpoint when milling a hole or trench—for example, when it is desired to make a hole or trench in the backside of the substrate of a flip-chip packaged IC. Optionally, a coarse trench may be milled in the thinned backside of the IC using either laser-stimulated chemical etching or ion-beam milling. A further small trench is milled using the focused ion beam. The endpoint for milling this small trench is determined precisely by monitoring the power supply leakage current of the IC induced by electron-hole pairs created in the milling process. This can be used to generate a precise in situ endpoint by modulating the beam at a reference frequency and then amplifying that frequency component in the leakage current using a narrow-band amplifier or lock-in amplifier. The precise in situ endpoint signal can be used for manual or automatic precise in situ endpoint detection. Different sputtered materials will give rise to different electron-hole pair signal levels, as shown in Fig. 1. This approach avoids or minimizes the unintentional damage or disturbance of the active diffusion regions
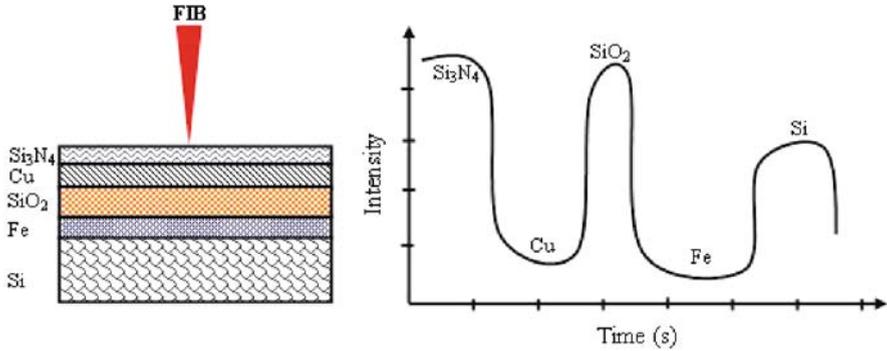
Fig. 1 Illustration of current signal for the endpoint detection when using FIB machines for milling in IC chip modification

in the IC by the ion beam. This allows normal operation of the IC after cutting the hole or trench. H.X. Qian et al. [3] reported their results by means of using the endpoint detection to detect and measure the thin film in real time, especially for the films with thickness less than the wavelength of light. Normally, they are too difficult to measure by conventional metrology methods. This is particularly useful when etching large aspect ratio holes or grooves, where a conventional endpoint detection-based secondary electron signal is inadequate and leads to large operator-induced uncertainties.

Figure 2 presents an optical micrograph of one of our examples of IC chip modification with joined and cut interconnect lines. For cutting interconnect lines, a large diameter hole is drilled across the line and filled with SiO_2 for insulating from the surrounding metal lines. For joining interconnects, a hole is drilled and filled with W. FIB machining can also be used for removing the protection layer on the top surface
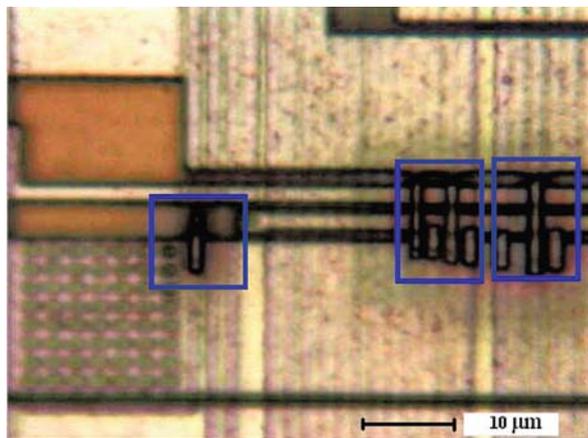
Fig. 2 Example of IC chip modification, a microdevice modification using FIB technology. Line cutting and joining were finished using FIB milling and deposition

Fig. 3 A packaged IC chip was etched selectively using gas assisted etching with the FIB machine. The chemical enchant gas was XeF_2. The protection layer of polymer on the top side was removed to expose the underlying metal interconnect lines
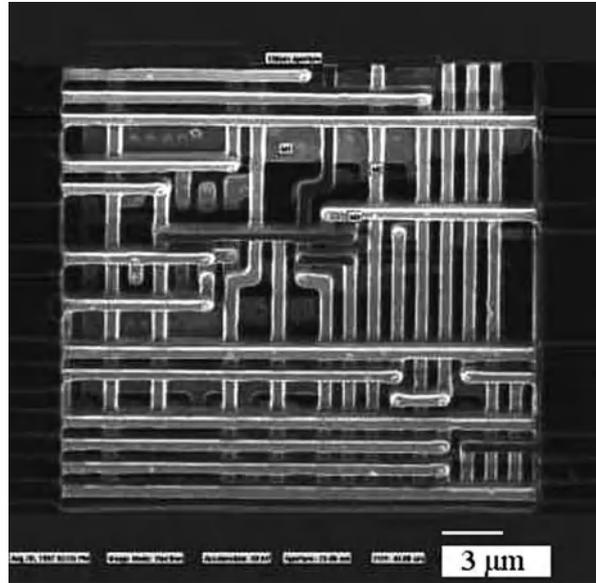
of the device using chemical gas-assisted etching, such as XeF_2 for polymers, as shown in Fig. 3.

References

1. P. Sudraud, G. Benassayag, and M. Bon, Microcircuit Engineering, Edited by R. Castagne and J. Perrocheau, p. 583, North Holland, 1987.
2. Y. Fu and N.K.A. Bryan, International Journal of Advanced Manufacturing Technology, 16(12), 877–880 (2000).
3. H.X. Qian, W. Zhou, X.M. Li, J.M. Miao, and L.E.N. Lim, Applied Surface Science, (2007) (in press).

Proton Beam Writing: A New 3D Nanolithographic Technique

Jeroen A. van Kan and Andrew A. Bettiol

1 Introduction

Current microelectronics production technologies are essentially two-dimensional (2D), and are well suited for the 2D topologies prevalent in microelectronics. As semiconductor devices are scaled down in size, and coupled with the integration of moving parts on a chip, there is expected to be a rising demand for smaller micro-electromechanical systems (MEMS) and nanoelectromechanical (NEMS) devices. High aspect ratio three-dimensional (3D) microstructures with nanometer details are also of growing interest for optoelectronic devices. Therefore it is essential to develop new lithographic techniques suitable for the production of high aspect ratio 3D micro- and nanocomponents. Proton-beam (p-beam) writing at the nanometer level is being developed at the Centre for Ion Beam Applications (CIBA), National University of Singapore, and has been shown to be a promising new 3D lithographic technique [1, 2]. In 2005 p-beam writing was incorporated in the Japanese road map for 3D nanofabrication [3]. Many other institutes have recently also started to look at p-beam writing, but mainly at the micron level. P-beam writing is a new technique that utilizes a focused beam of fast (MeV) protons written directly into a resist to produce a 3D latent image in a resist material (e.g., for sub-100 nm lithography, PMMA, SU-8, or HSQ). P-beam writing is the only technique that offers the capability of direct-write high aspect ratio nano- and microstructures. The secondary electrons induced by the primary proton beam have low energy and therefore limited range, resulting in minimal proximity effects. Low proximity effects coupled with the straight trajectory and high penetration of the proton beam enables the production of 3D micro- and nanostructures with well-defined smooth side walls. The current status of p-beam writing will be discussed. Recent tests have shown this technique capable of writing high aspect ratio walls up to 160 and details down to 20 nm in width with sub-3 nm edge smoothness [1, 4]. Computer simulations have

J.A. van Kan (✉)

Department of Physics, Centre for Ion Beam Applications (CIBA), National University of Singapore, Singapore 117542, Singapore
e-mail: phyjavk@nus.edu.sg

confirmed minimal proximity effects in p-beam writing [5]. These features make p-beam writing a direct write technique of high potential for the production of high aspect ratio structures with sub-100 nm detail in the lateral directions for rapid prototyping. In combination with electroplating, p-beam writing can fabricate precise 3D metallic moulds and stamps for batch and high-volume production, using either nanoimprinting or soft imprint lithography. Potential applications will be discussed in Boxes 10, 11, 12 and 13.

P-beam writing is a new technology with no commercial instruments available as yet. Technical and commercial development of p-beam writing have been hampered by the difficulties encountered in focusing MeV ions to sub-100 nm dimensions. These difficulties have recently been overcome [6] and the first prototype p-beam writer has recently been constructed at CIBA [7].

2 Design of a Proton Beam Writing System

2.1 P-Beam Exposure Station

In the early stage, p-beam writing was performed using a belt driven Van de Graaff accelerator (HVEC AN 2500). This type of accelerator in general exhibits poor energy stability, and consequently produces proton beams with wide intensity fluctuations (typically in the tens of Hz frequency range). Therefore, in order to achieve accurate exposure doses, a beam normalization procedure has to be developed: see Sect. 2.2. Initially p-beam writing was performed using the Oxford Microbeams OM2000 end station in a nonoptimized setup. Typical feature sizes that can be obtained with this kind of setup are down to the submicron level. More details about the setup used in CIBA in combination with the HVEC AN2500 Van de Graaff accelerator can be found in [8–11]. There are several other groups around the world active in p-beam writing with either a Van de Graaf accelerator or a Pelletron type of accelerator [12–16]. Some groups are doing masked proton beam lithography with MeV protons [17] or keV protons [18–20].

Recently several groups started to use cascade accelerators for p-beam writing [6, 21–24]. In Table 1 a summary is given of the activities of different groups in developing p-beam writing. Only with the cascade accelerators have people come close to the 100 nm level. Most of the systems used are not dedicated p-beam writing systems. In CIBA the first dedicated system for p-beam writing has been developed and has achieved 22 nm p-beam written features and spot sizes down to 35×75 nm^2.

In the rest of this chapter the CIBA system will be discussed, since it is currently the only system capable of routinely writing deep sub-100 nm features using MeV protons.

In 2000 a 3.5 MV HVEE Singletron^{TM} accelerator was installed in CIBA. P-beam writing with this accelerator can routinely achieve spot sizes at the sub-100 nm level because of its high brightness and increased beam stability. In addition a new nuclear nanoprobe facility has been developed at CIBA [6, 7]; see Fig. 1. This facility is the first of its kind dedicated to p-beam writing on a micro- as well as on a nanoscale.

Table 1 Current status and charge density requirements in PBW

Resist	Lab involved	Type	Beam charge density needed (nC/mm^2)	Smallest feature written
PMMA [44–46]	CIBA	Positive	80–150	20–30 nm
SU-8 [1, 44]	CIBA	Negative	30	60 nm
HSQ [4]	CIBA	Negative	30	22 nm
PMGI [44]	CIBA	Positive	150	1.5 μm
WL-7154 [33]	CIBA	Negative	4	800 nm
TiO_2 [33]	CIBA	Negative	8000	5 μm
Si [47]	CIBA	Negative	80,000	15 nm tip (implanted in channeling geometry)
DiaPlate 133 [48]	Le Locle, Switzerland	Negative	10	10 μm
ADEPR	Debrecen, Hungary [49]	Negative	125–238	5 μm
	CIBA [50]		25??	110 nm
Forturan [51]	Debrecen, Hungary	Positive	1	3 μm
PADC (CR-39) [51, 52]	Debrecen, Hungary	Positive	600	5 μm
ma-N 440 [53]	Leipzig, Germany	Negative	200	400 nm
GaAs [54]	Surrey, UK	Negative	100,000	12 μm
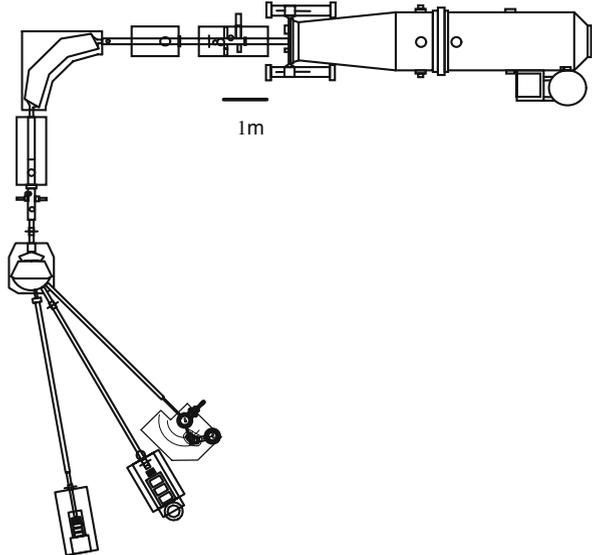
Fig. 1 Schematic diagram of the beam line facilities at CIBA. The p-beam writing facility is located under 10°; the 30° and 45° are the nuclear microscope facility and the high resolution RBS facility, respectively

In the CIBA p-beam facility protons from the accelerator are energy analyzed using a 90° magnet. Beam defining object slits are positioned one meter in front of the switcher magnet. Here we can set a rectangular object aperture, such that the transmitted beam through the aperture is brought into the 10° beamline and demagnified with the Oxford Microbeams high demagnification lenses (OM52) in a high excitation triplet configuration. This lens system operates at an object distance of 7 m and a reduced image distance of 70 mm, resulting in enhanced system demagnifications (228×60 in the X and Y directions, respectively).

To prevent deposition in the sample of any unwanted dose we use a beam blanking system located directly behind the object slits, where the beam is deflected out of the normal beam path using the field generated between a set of electrostatic plates. Typical switching time is $0.5\,\mu s$, fast enough for the scanning speeds used in p-beam writing experiments.

The end station is supported by an optical table for increased stability. The control electronics and roughing pumps are kept several meters away from the end station in order to reduce beam spot broadening due to stray AC electromagnetic fields [25]. Because of the short working distance between the lens and the target, the optical microscope installed to check the position of the target has a limited magnification due to adverse geometrical considerations.

The resist sample is mounted on a computer-controlled Burleigh Inchworm EXFO XYZ stage which has a travel of 25 mm for all axes with a 20 nm closed loop resolution. The system has been designed to be compatible with Si wafers up to 6″. One of the critical features of a stage for p-beam writing is the fact that it doesn't introduce any AC magnetic field.

During exposures the beam is scanned over the resist using a set of electromagnetic scan coils, located directly in front of the quadruple lens system. In this way scan fields up to $0.5 \times 0.5\,mm^2$ can be achieved. Up to now only magnetic scanning has been used for p-beam writing. Magnetic scanning systems have a relatively long settling time because of the magnetic scan coils, resulting in a relatively slow writing speed. Tests show that with a prototype electrostatic scanning system, proton beam writing speeds can be increased by a factor of 15 compared with magnetic scanning [26]. Here the writing speed was limited by the speed of the electrostatic scanning high voltage power supply. In addition, we have also been utilising stage scanning, which allows us to scan the beam over the full $25 \times 25\,mm^2$ range of the XYZ stage.

To allow fast focusing of the proton beam, provision is made to use signals such as secondary electron emission and ion-induced photon emission, which typically have a much higher yield per proton compared to the number of nuclear backscattered events per proton

2.2 Scanning System

Scanning a focused MeV proton beam usually involves the use of either a transimpedance amplifier (magnetic scanning, current output) or a high voltage amplifier (electrostatic scanning, voltage output). Both types of amplifiers take a bipolar

or unipolar voltage signal as an input, which is usually less than 10 volts. This signal is used to drive a set of magnetic scan coils or electrostatic scan plates. In order to scan a proton beam, one needs to generate a precise two-dimensional arbitrary voltage waveform. The starting point for generating such an arbitrary waveform is to create a scan pattern in the form of an image file on a computer. The irradiation can be represented by, for example, a black and white bitmap file or an AutoCAD dxf file. If a bitmap file is used, then the black pixels represent the irradiated regions and the white pixels signify regions in the scan field that are not irradiated. The points in the scan pattern file need to be converted into a list of x and y coordinates corresponding to digitized voltage values that can be generated by a digital-to-analog converter (DAC) on a computer card. The voltage range is divided into 2^n pixels, where n is an integer typically between 8 and 16. This gives scan patterns with a digital resolution between 256×256 and 65536×65536. The actual pixel resolution can be determined by dividing the scan size (in μm) by the digital resolution. In order to generate the x and y waveforms, two DACs on a multi I/O data acquisition card are used. A third DAC is also used to generate a blanking level which deflects the beam when it is in regions that should not be irradiated. The three DACs are updated simultaneously at a minimum rate that depends on the hardware being used. The National Instruments NI 6731 card that is used at CIBA allows one to clock the three DACs at a minimum rate of $1\,\mu$s with a maximum digital resolution of 16 bits.

The main software package used for proton beam writing at various laboratories around the world is called Ionscan [27]. This software was specifically developed to enable one to scan an arbitrary pattern while simultaneously monitoring a normalization signal. This normalization signal is in the form of a TTL pulse that originates from any detector that measures the proton solid interaction, and is proportional to the fluence delivered to the sample. The types of signals that have been used for normalization include Rutherford backscattering spectrometry (RBS), particle induced X-ray emission (PIXE), charge, electron emission, and light emission. In order for the Ionscan software to be able to deliver the correct particle fluence, the sample must first be calibrated. The most commonly used method for fluence normalization and calibration is RBS. To use this calibration method, a particle detector with a known surface area and in a known geometry is required. This detector is usually a silicon surface barrier detector. At the start of an experiment, an RBS spectrum is collected from the sample with sufficient statistics for the data to be fitted using an analysis software package like SIMNRA [28]. From the fit, the number of ions * steradian in the spectrum and the total number of normalization counts are extracted and entered into the appropriate Ionscan dialog window. The Ionscan software uses this data, along with other information like the scan size and the desired fluence, to calculate a number called the *normalization number*. This is the number of counts required in order to expose the sample with the desired fluence measured by the detector used for normalization. Once the scan file has been read, and the sample calibration details entered into Ionscan, one of four different normalization modes, preset, pixel, shape, and figure, can be used to deliver the desired fluence to the sample. If preset normalization is chosen, the user can enter the number of times that the

pattern is repeatedly scanned. If this normalization mode is chosen, the user needs to calculate the total time required for the exposure based on the beam current (detector count rate) and the update time between the irradiated pixels. This method produces excellent results if the beam has a high degree of stability. For pixel normalization, the beam remains at each pixel until the required fluence (counts) is measured by the normalization detector. Pixel normalization is potentially the most accurate normalization mode; however, if the number of counts required for normalization per pixel is low, then this method is prone to substantial errors. Typically if RBS is used as the normalization signal for low current proton beam writing at submicron dimensions, then the number of counts per pixel may be less than 1. In this case one of the other normalization methods is preferable. *In the CIBA system for writing at the 100 nm level and below, direct proton counting (using a photodiode) can be utilized to good effect in combination with the preset scan normalization mode.* Shape normalization involves repeatedly scanning the beam over a shape until the desired fluence is reached before moving to the next shape. A shape in proton beam writing is defined as a group of adjacent pixels that are separated from the next group of pixels by a path where the beam is blanked. This normalization method decreases the number of times the beam is scanned over regions of the sample where the beam is blanked. The final normalization method available is called figure scanning. This method scans the whole pattern, including regions where the beam is blanked, until the desired fluence is reached. This method is essentially an automated version of preset scanning. The fluence normalization method chosen for a particular experiment will depend on the sample to be irradiated, the size of the structure to be irradiated, the beam spot resolution, and the desired fluence.

3 Experimental Procedures in Proton Beam Writing

3.1 Focusing Issues at the Sub-100 nm Level

The CIBA p-beam system is able to focus proton beams down to a $35 \times 75 \, nm^2$ spot size, which can be used for maskless direct write lithography [6]. In the first tests with this system the focusing power of the new lenses was tested. A beam of 1 MeV protons was scanned over a $2 \, \mu m$ thick freestanding X-ray mask which contained $1 \times 1 \, \mu m^2$ holes. The data was collected using a scanning transmission ion microscopy (STIM) signal with an energy window centred on the incident beam energy; see Fig. 2. Deconvolution of line scans perpendicular to the edges of a $1 \, \mu m$ hole, which are assumed to be sharp and vertical, indicate a beam spot size of $35 \times 75 \, nm^2$. Scanning electron micrograph (SEM) analysis of the X-ray mask used to deconvolute the beam spot size indicates that the side walls of the holes are not perfect, but have a natural width profile of about 30 nm. This represents a limiting factor in the determination of the beam spot size. To improve the beam focusing below the 100 nm level, a p-beam written resolution standard has been developed.
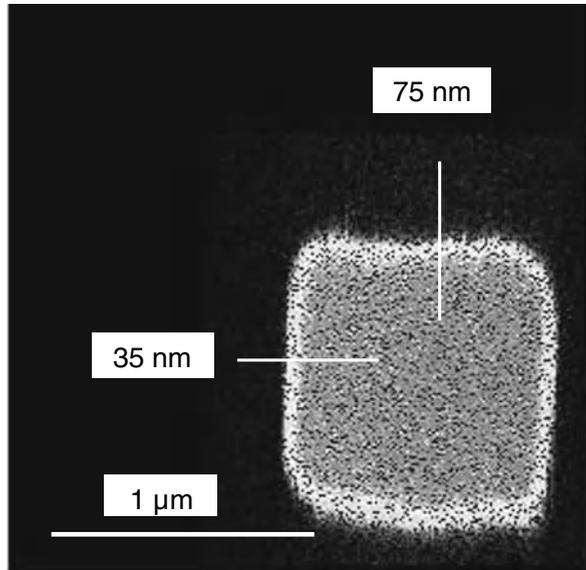
Besides smooth sidewalls, it is important to have perfect sidewall verticality in a resolution standard. SRIM [29] calculations show that a parallel incoming proton beam will spread less than 8.0 nm (90% of the beam) after penetrating $2\,\mu m$ in PMMA. Secondary electron excitation calculations [30] show that 90% of the energy will be deposited within 3.0 nm of the proton track using a 2 MeV proton beam in $2\,\mu m$ thick PMMA; a typical beam divergence on target is $0.1°$. Taking these facts into consideration we can expect a sidewall verticality of about $89.6°$ if we produce a $2\,\mu m$ thick resolution standard. This thickness is a good starting point to make a resolution standard, considering that the noncommercial X-ray mask previously used as a sub-100 nm resolution standard has a sidewall verticality of $89.1°$ [6, 7]. A thickness of $2\,\mu m$ is a reasonable compromise between stability of a freestanding Ni grid and the sidewall width projection, which we expect to be about 12 nm, compared to 30 nm in the previously used X-ray mask. The resolution standard was written in a $2.5\,\mu m$ thick layer of PMMA which was subsequently Ni plated and released through a sacrificial layer etching [31, 32].

A 2D map was recorded employing a 2 MeV H_2^+ beam with a current of 15,000 protons per second over an edge in the Ni grid. A horizontal line scan was extracted and showed a beam size of 29.2 nm FWHM. The data was collected using a scanning transmission ion microscopy (STIM) signal with an energy window centred on the incident beam energy. To accumulate enough statistics, 6 scans were performed over the same area; see Fig. 3 [31].

Nanomachining at high accuracy requires the development of a focusing protocol, which ideally should include computer software for beam spot optimization, and high quality resolution standards. Both of these are currently being developed at CIBA.
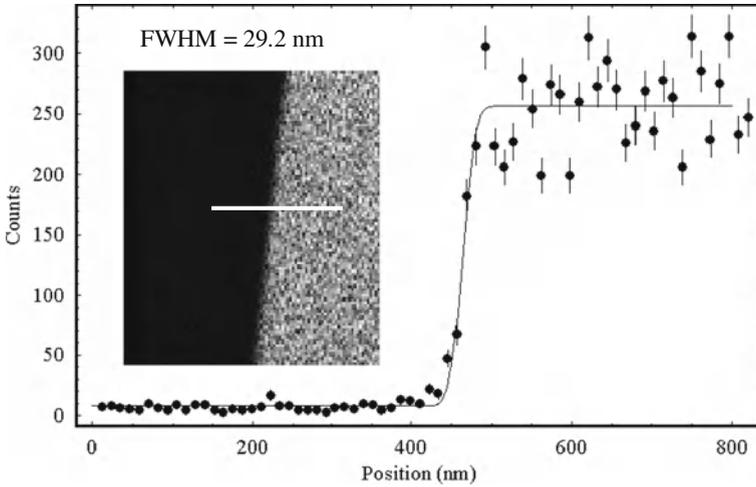
Fig. 3 STIM map of an edge of the Ni grid (inset) and extracted line scan reproduced from [31]

3.2 Proton-Resist Interactions

In p-beam writing the path of a high energy (MeV) proton in material is dependent on the interaction with the electrons and nuclei in the resist material. The probability that a proton interacts with an electron is a few orders of magnitude higher than for nuclear scattering in the first 50% of its trajectory. Therefore only proton–electron interactions need to be considered in the first part of the proton trajectory. Proton–electron interactions hardly change the trajectory of a proton because of the mass ratio $(m_p/m_e \sim 1800)$, which implies that the path of a proton hardly deviates from a straight line. Since the energy transfer in these collisions is rather small, peaked at around 100 eV, many collisions will occur before a proton comes to rest. Proton trajectories can therefore be accurately simulated by means of Monte Carlo calculations—for example, using the computer code SRIM [29]. In electron beam writing as well as in p-beam writing the induced secondary electrons are either chain scissioning or cross linking the molecules in positive and negative resist, respectively. Therefore, resists which are useful as electron-beam resists are potential candidates for p-beam writing [33]. The physical characteristics of ion/electron interactions make p-beam writing a predictable and an extremely powerful lithographic technique with the following key features:

(a) *Protons have a relatively long and well-defined range in resist materials*. The penetration depth depends on the proton energy; e.g., the penetration of a 1.0 MeV proton in PMMA is 20 μm, whereas a 3.5 MeV proton will penetrate 160 μm [29]. This feature allows the production of slots and holes of well-defined depth, and the creation of multilevel structures in one resist layer [11, 34]. By exposing the negative resist SU-8 to protons at different energies, novel structures can be produced in

one layer of resist. These structures include buried microchannels, cantilevers, etc. [11, 35].

(b) *The proton beam travels in a straight line*, with very little small angle scattering except at the end of range. This allows the production of structurally accurate high aspect ratio structures. The confinement of the secondary electrons around the proton track coupled with an even energy deposition along the path of the proton beam generates smooth sidewalls. Our calculations show that lithography with MeV protons is potentially capable of producing high quality nano- and microstructures with 1 nm smoothness and high aspect ratio

(c) *The proton beam has a relatively even dose distribution with penetration depth*. Monte Carlo calculations using SRIM [29] show that the energy deposition increases slowly with depth, and increases rapidly only at the end of range, where energy loss due to proton/nuclear collisions increase. This feature ensures a relatively even exposure distribution with depth. This contrasts with 3D lithography using penetrating electromagnetic radiation (e.g., X-rays), which exhibits an exponential dose distribution with depth.

(d) *Reduced proximity effects*. One big potential advantage of proton lithography is the virtual absence of high energy secondary electrons which can give rise to unwanted exposure of resist in e-beam lithography. In e-beam writing, much of the energy is dissipated in the form of secondary electrons with an energy of 2–50 eV, with a small but significant fraction of the secondary electrons having significant energies which can contribute to the proximity effect in the range of a few tenths of microns [36].

3.3 Sub-100 nm Resist Materials for P-Beam Writing

Although several resists have been tested for compatibility with p-beam writing [33, 37], only polymethylmethacrylate (PMMA A11 950k), SU-8 (a chemically amplified, epoxy-based resist) and hydrogen silsesquioxane (HSQ) resist were found to be suitable for writing at the sub-100 nm level. PMMA and SU-8 are obtained from MicroChem Corp. (MCC); HSQ is obtained from Dow Corning. Earlier tests have shown that other resists (e.g., PMGI, WL-7154, TiO_2, etc.) can also be machined with a reduced resolution [33]; tests using Novolac resists have been disappointing, since these resists show positive as well as negative resist behavior under p-beam writing [37].

Typical samples used in p-beam writing consist of a flat substrate (e.g., Si or glass) coated with a resist layer. Prior to resist coating the samples are cleaned in Pirhana etch solution (H_2SO_4 : H_2O_2 3 : 1) followed by de-ionized water rinse and a dry bake (10 min at 200 °C) to remove any moisture. Next the samples are spin-coated with resist followed by a baking step to remove the solvent, 180 °C, 95 °C, and 150 °C are used to bake PMMA, SU-8, and HSQ resist, respectively. Typical baking time is several minutes, the exact baking time depends on the thickness and type of resist.

3.3.1 PMMA Resist

In the exposure of PMMA, the protons cause chain scissioning of the polymer chains. The resulting damaged resist, consisting of molecular chains with lower molecular weight, are then selectively removed using either GG developer at 30 °C [38] or a less viscous developer for nanostructures, comprising isopropyl alcohol (IPA) and DI water (7:3). The optimum dose for PMMA resist using 2 MeV protons is either 90 or 150 nC/mm^2 for the two different developers respectively. PMMA therefore is a positive resist under proton irradiation. In Fig. 4 we see 50 nm wide lines written with a 2 MeV proton beam focused down to a 100 nm spot size, in 350 nm thick PMMA [1]. For the fabrication of these precise structures there was no special exposure strategy, i.e., the lines were exposed using uniform proton beam dose distribution, which suggests no proximity effect is present in this experiment.

Fig. 4 SEM image of parallel lines written in a 350 nm thick PMMA layer. The structure was written with a focused 2 MeV proton beam. The photo indicates a wall width of 50 nm. Reproduced from [1]
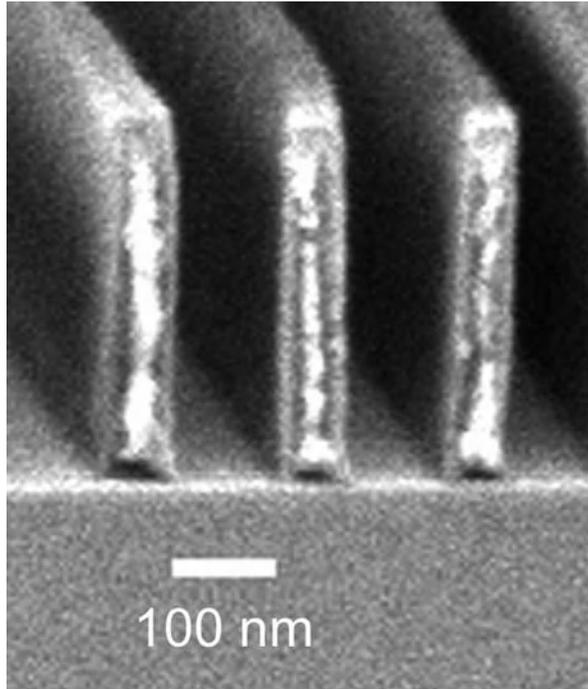
3.3.2 SU-8 Resist

During irradiation the proton-induced electrons will cross link the SU-8 resist; the required dose for full crosslinking using 2 MeV protons is 30 nC/mm^2. After irradiation the resist will be developed. To selectively remove the unexposed areas, SU-8 is developed using the standard developer from MCC at room temperature [37],
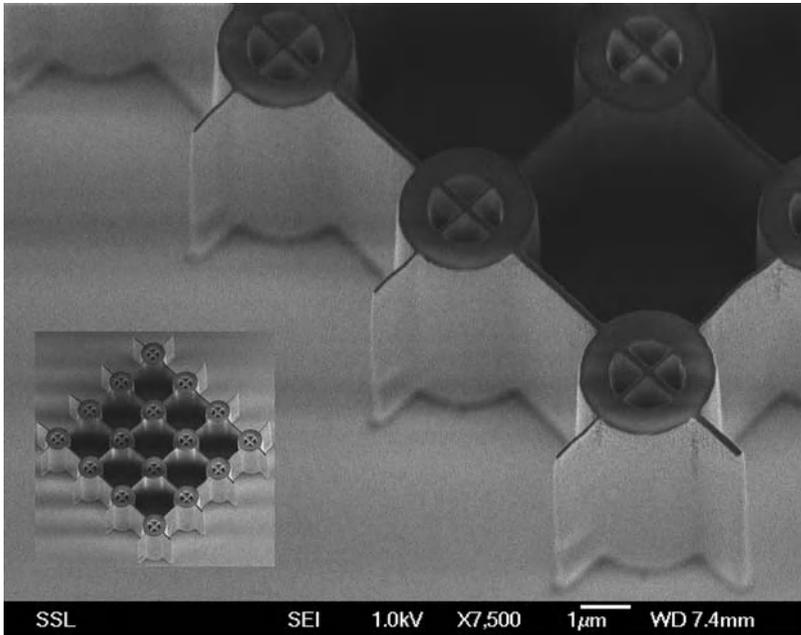
Fig. 5 High aspect ratio test structures fabricated using p-beam writing in SU-8 negative resist, showing 60 nm wall structures 10 μm deep. Reproduced from [39]

making SU-8 a negative resist under proton beam exposure. Note that in the case of SU-8, no post-exposure bake is required. Tests with post-exposure bake have shown that structures with fine details ($<$ 500 nm) are destroyed during post-exposure bake.

In Fig. 5 we see an array of micropillars connected via 60 nm walls in a 10 μm thick SU-8 layer [39]. SEM analysis indicates near vertical sidewalls. This corresponds to an extremely high aspect ratio in SU-8 of 160 (height : width). We believe this to be a state-of-the-art performance in SU-8; only Bogdanov and Peredkov [40] have reported similar aspect ratios for SU-8 structures, but with structures of 4 μm width or more.

3.3.3 HSQ Resist

HSQ is a relatively new p-beam resist. Typically HSQ can be applied in layers up to 2 μm. A silicon wafer was coated with an 850 nm thick layer of HSQ (Fox-17, Dow Corning) by spin coating the silicon wafer for 30 s at 3000 rpm. The wafer was pre-baked for 120 s at 150 °C after spin coating. After proton exposure the sample was developed in a 2.38% tetramethyl ammonium hydroxide (TMAH) solution for 60 s followed by a DI water rinse. The contrast curve for the 850 nm layer was measured, and a contrast of 3.2 was found for p-beam writing [4]. Here the contrast

Fig. 6 SEM images of a
22 nm wide line written with
a 2 MeV H_2^+ beam using a 3
pixel wide exposure pattern
in 850 nm thick HSQ.
Reproduced from [4]
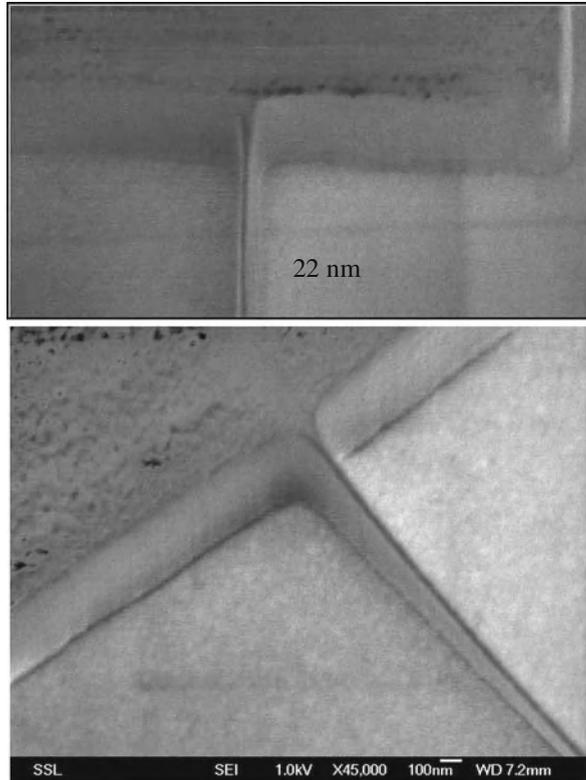
is defined as $\gamma = 1/[\log (D_f) - \log(D_i)]$, where D_f is the dose at which the resist
is fully insoluble and D_i the dose where the resist becomes insoluble. Similar con-
trast values have been reported for e-beam writing in HSQ [41, 42]. We define the
sensitivity as the point where the layer is fully insoluble and reaches the maximum
thickness; and for protons on HSQ we have measured a sensitivity of 30 nC/mm^2,
similar to the sensitivity found for SU-8 exposure with protons [8]. This definition
for sensitivity is used because a lower exposure dose was found to result in weaker
HSQ structures. Wall structures were written with 2 MeV H_2^+ in this HSQ layer. A
line was scanned three times in a total exposure pattern of 1.4 μm^2; see Fig. 6. Here
a fluence of 1.2×10^6 protons was used to expose the pattern. After development,
a line width of 22 nm was observed. This corresponds to an aspect ratio of 39:1.
The wall is slightly tilted because of capillary forces during development. Down
to the 20 nm level we have shown that by exposing HSQ with a sufficient proton
dose, the walls remain standing without the use of supercritical drying necessary for
successful development of e-beam written HSQ structures [43].

References

1. J.A. van Kan, A.A. Bettiol and F. Watt, APL **83** (2003) 1629–1631.
2. G. Chin, Science **301** (2003) 1291–1292.
3. http://www.nbci.jp/file/051025-4.pdf
4. J.A. van Kan, A.A. Bettiol and F. Watt, Nano Lett. **6** (2006), 579.
5. CNB Udalagama, AA Bettiol and F Watt, Nucl. Instrum. Methods **B260** (2007) 384–389.
6. F. Watt, J.A. van Kan, I. Rajta, A.A. Bettiol, T.F. Choo, M.B.H. Breese and T. Osipowicz, Nucl. Instrum. Methods **B210** (2003) 14–20.
7. J.A. van Kan, A.A. Bettiol and F. Watt, Mater. Res. Soc. Symp. Proc. 777 (2003) T2.1.1.
8. J.L. Sanchez, J.A. van Kan, T. Osipowicz, S.V. Springham, and F. Watt, Nucl. Instrum. Methods **B136–B138** (1998) 385.
9. J.A. van Kan, J.L. Sanchez, B. Xu, T. Osipowicz and F. Watt, Nucl. Instrum. Methods, **B148** (1999) 1085.
10. J.A. van Kan, A.A. Bettiol, B.S. Wee, T.C. Sum, S.M. Tang and F. Watt, Sens. Actuat. **A92** (2001) 370–374.
11. F. Watt, J.A. van Kan and T. Osipowicz, MRS Bull. **25** (2000) 33.
12. V. Auzelyte, M. Elfman, P. Kristiansson, J. Pallon, M. Wegdén, C. Nilsson, K. Malmqvist, B.L. Doyle, P. Rossi, S.J. Hearne, P.P. Provencio and A.J. Antolak, Nucl. Instrum. Methods **B242** (2006) 253.
13. Y. Gonin, F. Munnik, F. Benninger, F. Dias and S. Mikhaïlov, J. Vac. Sci. Technol. **B22** (2004) 1982.
14. I. Rajta, E. Baradács, M. Chatzichristidi, E.S. Valamontes, I. Uzonyi and I. Raptis, Nucl. Instrum. Methods **B231** (2005) 423.
15. B. Rout, M. Kamal, A.D. Dymnikovm, D.P. Zachry and G.A. Glass, Nucl. Instrum. Methods, **B260** (2007) 366.
16. D.H. Morse and A.J. Antolak, *17th Conference on Application of Accelerators in Research and Industry* (2003) 625.
17. U. Weidenmüller, J. Meijer, A. Stephan, H.H. Bukow and C. Rolfs, Microelectron. Eng. **46** (1999) 489.
18. I. Adesida, Nucl. Instrum. Methods **B7/8** (1985) 923.
19. G. Stengl, R. Kaitna, H. Löschner, P. Wolf and R. Sacher, J. Vac. Sci. Technol. **16** (1979) 1883.
20. G. Stengl, R. Kaitna, H. Löschner, R. Rieder, P. Wolf and R. Sacher, J. Vac. Sci.Technol. **19** (1981) 1164.
21. F. Menzel, D. Spemann, J. Lenzner, J. Vogt and T. Butz, Nucl. Instrum. Methods **B231** (2005) 372.
22. L. van IJzendoorn and M. de Voigt, Nederlands tijdschrift voor Natuurkunde **7** (2004) 238.
23. P. Mistry, I. Gomez-Morilla, G.W. Grime, R.P. Webb, R. Gwilliam, C. Jeynes, A. Cansell, M. Merchant, R. Smith and K.J. Kirkby, Nucl. Instrum. Methods **B260** (2007) 437.
24. J. Simčič, P. Pelicon, Z. Rupnik, M. Mihelič, A. Razpet, D. Jenko and M. Maček, Nucl. Instrum. Methods **B241** (2005) 479.
25. F. Watt, T.F. Choo, K.K. Lee, T. Osipowicz, I. Orlic and S.M. Tang, Nucl. Instrum. Methods **B104** (1995) 101.
26. J.A. van Kan, A.A. Bettiol, K. Ansari, S. Peige and F. Watt, Proceedings IEEE MEMS (2004), 673–676.
27. A.A. Bettiol, J.A. van Kan, T.C. Sum and F. Watt, Nucl. Instrum. Methods, **B181** (2001) 49–53.
28. W. Eckstein and M. Mayer, Nucl. Instrum. Methods **B153** (1999) 337.
29. J. P. Biersack and L. G. Haggmark, Nucl. Instrum. Methods **174** (1980) 257. Recently updated, the package and its documentation are available at http://www.research.ibm.com/ionbeams/SRIM/SRIMINTR.HTM
30. M.P.R. Waligorski, R.N. Hamm and R. Katz, Nucl. Tracks Radiat. Meas. **11** (1986) 309–319.

31. J.A. van Kan, P.G. Shao, P. Molter, M. Saumer, A.A. Bettiol, T. Osipowicz and F. Watt, Nucl. Instrum. Methods **B231** (2005) 170.
32. F. Zhang, J. A. van Kan, S.Y. Chiam and F. Watt, Nucl. Instrum. Methods **B260** (2007) 474.
33. J.A. van Kan, A.A. Bettiol, S.Y. Chiam, M.S.M. Saifullah, K.R.V. Subramanian, M.E. Welland and F. Watt, Nucl. Instrum. Methods **B260** (2007) 460.
34. J.A. van Kan, J.L. Sanchez, T. Osipowicz and F. Watt, Microsyst. Technol. **6** (2000), 82–85.
35. J.A. van Kan, A.A. Bettiol, K. Ansari and F. Watt, Proc. SPIE **4343** (2001) 466–472.
36. P. Rai-Choudhury, *Handbook of Microlithography, Micromachining and Microfabrication Volume 1: Microlithography* (SPIE, Washington, 1997), Chapter 2 Electron Beam Lithography M.A. Mc Cord, & M.J. Rooks.
37. J.A. van Kan, J.L. Sanchez, B. Xu, T. Osipowicz and F. Watt, Nucl. Instrum. Methods **B158** (1999) 179–184.
38. S.V. Springham, T. Osipowicz, J.L. Sanchez, L.H. Gan and F. Watt, Nucl. Instrum. Methods **B130** (1997) 155–159.
39. F. Watt, A.A. Bettiol, J.A. van Kan, E.J. Teo and M.B.H. Breese, Int. J. Nanosci. **4** (2005) 269.
40. A.L. Bogdanov and S.S. Peredkov, http://www.maxlab.lu.se/beamlines/bld811
41. F.C.M.J.M. van Delft, J.P. Weterings, A.K. van Langen-Suurling and H. Romijn, J. Vac. Sci. Technol. **B18** (2000) 3419.
42. F.C.M.J.M. van Delft, J. Vac. Sci. Technol. **B20** (2002), 2932.
43. H. Namatsu, J. Vac. Sci. Technol. **B19** (2001), 2709.
44. J.A. van Kan, J.L. Sanchez, B. Xu, T. Osipowicz and F. Watt, Nucl. Instrum. Methods **B158** (1999) 179.
45. J.A. van Kan, K. Ansari, P. Shao, A.A. Bettiol and F. Watt, *Book of Abstracts*, NNT 2004, Vienna.
46. M.B.H. Breese, G.W. Grime and F. Watt, Nucl. Instrum. Methods **B77** (1993) 169.
47. E.J. Teo, M.B.H. Breese, E.P. Tavernier, A.A. Bettiol, F. Watt, M.H. Liu and D.J. Blackwood, Appl. Phys. Lett. **84** (2004) 3202.
48. Y. Gonin, F. Munnik, F. Benninger, F. Dias and S. Mikhaïlov, J. Vac. Sci. Technol. **B22** (2004) 1982.
49. I. Rajta, E. Baradács, M. Chatzichristidi, E.S. Valamontes, I. Uzonyi and I. Raptis, Nucl. Instrum. Methods **B231** (2005) 423.
50. M. Chatzichristidi, E. Valamontes, P. Argitis, I. Raptis, J.A. van Kan, F. Zhang and F. Watt, in print Microelectron. Eng., doi:10.1016/j.mee.2007.12.005
51. I. Rajta, I. Gómez-Morilla, M.H. Abraham, Á.Z. Kiss, Nucl. Instrum. Methods **B210** (2003) 260.
52. I. Rajta, E. Baradács, A.A. Bettiol, I. Csige, K. Tőkési, L. Budai and Á.Z. Kiss, Nucl. Instrum. Methods **B231** (2005) 384.
53. F. Menzel, D. Spemann, S. Petriconi, J. Lenzner and T. Butz, Nucl. Instrum. Methods **B260** (2007) 419.
54. P. Mistry, I. Gomez-Morilla, G.W. Grime, R.P. Webb, R. Gwilliam, C. Jeynes, A. Cansell, M. Merchant, R. Smith and K.J. Kirkby, Nucl. Instrum. Methods **B260** (2007) 437.

Box 10: Proton Beam Writing of Optical Structures

Andrew A. Bettiol

Proton beam writing, being a direct write lithographic technique, is well suited to rapid prototyping of micro- and nanophotonic devices [1]. The technique can be applied to a wide range of materials, including polymers [2], semiconductors [3], and glasses [4, 5]. Furthermore, the minimum feature size that can be routinely obtained is well matched to current photonic device technology (100 nm–1 μm resolution). We will now discuss in more detail how proton beam writing can be used to fabricate some common optical components that can be used to make more complicated photonic devices.

Optical waveguides are one of the basic building blocks of many microphotonic devices, including amplifiers, optical switches, ring resonators, and various types of interferometers that can be used for applications in biological and chemical sensing. There are two types of waveguiding structures that can be fabricated using proton beam writing. The first type of structure involves direct micromachining of the high refractive index core, followed by the coating of a lower refractive index cladding layer. Using this fabrication procedure, the core size and refractive index contrast of the waveguide can be easily controlled by choosing an appropriate polymer combination and spin coating the desired cladding thickness. This procedure is therefore useful for making both single mode and multimode waveguides in polymers. An example of a y-branch fabricated in a 5 μm layer of SU-8 spin coated on a Pyrex^{TM} substrate is shown in Fig. 1. The cladding layer chosen for this application was NOA-88, a UV curable polymer that has a refractive index of 1.555 at a wavelength of 632.8 nm [2]. The refractive index of the SU-8 core (1.596 at 632.8 nm) and the Pyrex^{TM} substrate (1.470 at 632.8 nm) were measured using a Metricon prism coupler system. The propagation loss of these waveguides is about 0.19 dB/cm at 632.8 nm. This was determined by measuring the intensity of the scattered light along the whole length of the waveguide using a CCD camera. The extremely low loss is due to the smoothness of the sidewalls of the fabricated structure. The

A.A. Bettiol (✉)

Department of Physics, Centre for Ion Beam Applications, National University of Singapore, Singapore 117542, Singapore

e-mail: phybaa@nus.edu.sg

Fig. 1 Y branch waveguide fabricated in a 5 μm layer of SU-8 spin coated on a Pyrex^TM glass substrate. (**a**) SEM image of the Y-branch before a cladding of NOA-88 adhesive has been applied, (**b**) SEM image of the end face after the cladding layer has been applied, (**c**) close-up of one of the waveguide cores
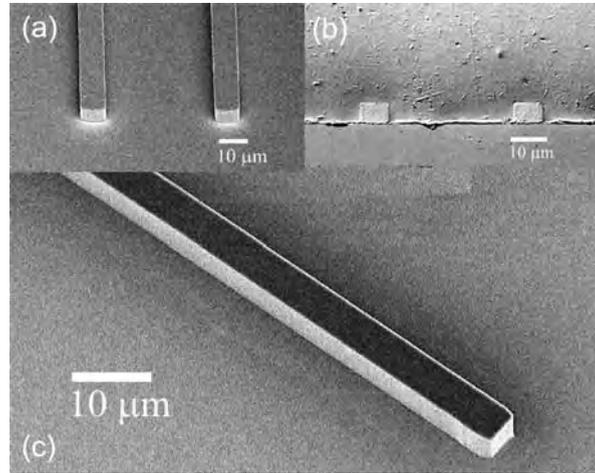
typical sidewall root mean square roughness of the SU-8 waveguides was measured using an AFM and shown to be approximately 4 nm [6]. This value is similar to those obtained in semiconductor waveguide structures etched using an inductively coupled plasma reactive ion etching [7].

Waveguides can also be fabricated in a single step if proton beam writing is used to modify the refractive index of the target material. This fabrication technique relies on the fact that the energetic mega-electron volt (MeV) ion impinging on a sample will lose most of its energy in a region adjacent to the end of range. This energy loss process is due either to the interaction of the energetic ion with target electrons (electronic energy loss), or to interaction with target nuclei (nuclear energy loss) resulting in atomic displacements. In some materials (e.g., fused silica), the ion beam-induced damage can cause an increase in density at the end of range, giving rise to a localized increase in refractive index. In order to achieve weak guiding of light, a refractive index contrast of about 10^{-4}–10^{-3} is usually sufficient, depending on the material being used. This refractive index contrast corresponds to a proton fluence of 10^{14}–10^{15} protons/cm^2 for fused silica [8] and 10^{11}–10^{14} protons/cm^2 for polymers [9]. An example of a waveguide fabricated in Foturan^TM glass using proton beam writing is shown in Fig. 2. Waveguiding was achieved in this material for fluences between 10^{14} and 10^{16} protons/cm^2. A refractive index increase at the end of range of 1.6×10^{-3} was measured for a fluence of 10^{16} protons/cm^2 [5]. The proton beam writing technique has also been applied to the fabrication of buried channel waveguides in phosphate glass that has been co-doped with Er^{3+}–Yb^{3+}. The fabricated waveguides have been used as optical amplifiers with a maximum net gain of 1.72 dB/cm measured at 1.534 μm [10]. A review of the optical effects of ion beam modification and the various materials used for waveguide formation can be found in [11].

The fabrication of surface relief gratings using proton beam writing in polymer has some unique processing challenges. Short period gratings that operate in the

Fig. 2 Ion implantation type
waveguide fabricated in
Foturan™ photosensitive
glass. (a) Mode profile
recorded for coupling with a
632.8 nm Helium Neon laser.
(b) Differential interference
contract optical image
showing the end of range of
the waveguide 45 μm below
the surface
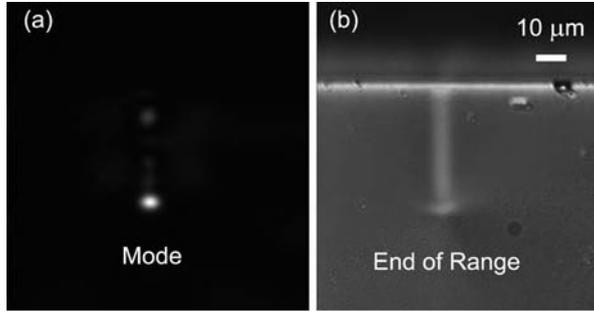
visible to near infrared wavelengths typically need to have a periodicity between
100 nm and a few microns. This places some stringent requirements on the accuracy
with which one knows the proton beam dimensions, since the minimum feature size
can be sub-100 nm. High density structures like gratings require precise control of
beam conditions like fluence and scanning signal noise in order to correctly expose
the desired structures and to avoid any unevenness. Furthermore, a good knowledge
of the development conditions is required, since it is almost impossible to observe
under an optical microscope if the structures are fully developed.
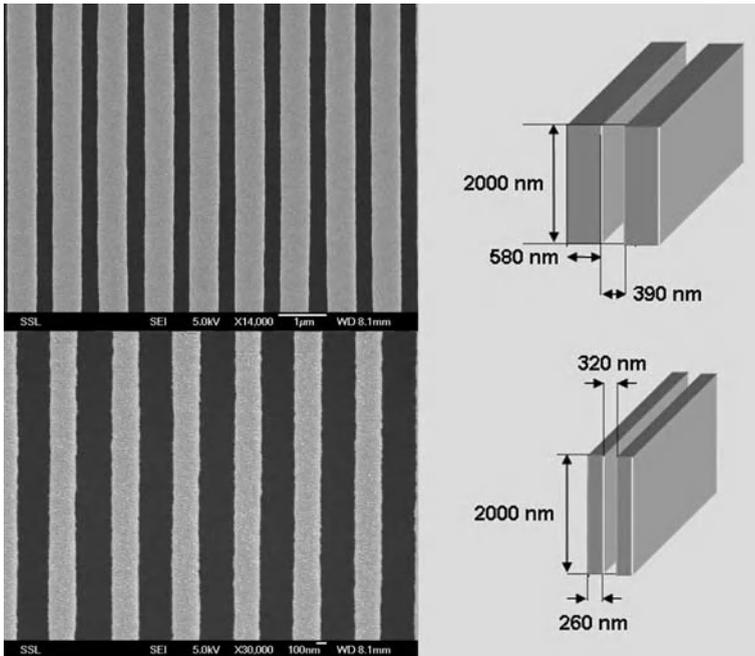
Fig. 3 Two gratings fabricated using proton beam writing in a layer of 2 μm PMMA resist

An example of two gratings fabricated using proton beam writing is shown in Fig. 3. The gratings were fabricated in a layer of $2\,\mu m$ poly-methylmethacrylate (PMMA) spin coated on a Si wafer with a Cr (20 nm)/Au (200 nm) seed layer. The whole grating structure is $100\,\mu m$ in length with a width of $30\,\mu m$. The metallic seed layer in these samples is typically used for subsequent electroplating, and it also assists in the adhesion of the PMMA to the silicon substrate. The extremely low line edge roughness can be easily observed in these images, indicating that higher density lines and spaces are possible with the proton beam writing technique. For both of these grating structures, the beam was repeatedly scanned for 6–10 loops in order to improve the line edge roughness. Fabrication of long period gratings is also possible using proton beam writing. von Bibra et al. [12] reported the use of proton beam writing for the fabrication of long period Bragg gratings in nonphotosensitive optical fiber. In order to fabricate the Bragg gratings, a proton energy of 2.4 MeV was chosen so that the end of range exactly corresponded to the depth at which the fiber core was located.

References

1. F. Watt, M. B. H. Breese, A. A. Bettiol, and J. A. van Kan: Mater. Today **30**, 20 (2007)
2. T. C. Sum, A. A. Bettiol, J. A. van Kan, F. Watt, et al: Appl. Phys. Lett. **83** 1707 (2003)
3. E. J. Teo, A. A. Bettiol, M. B. H. Breese, et al: Opt. Exp. **16** 573 (2008)
4. A. Roberts and M. L. von Bibra: J. Light. Technol. **14** 2554 (1996)
5. A. A. Bettiol, S. Venugopal Rao, E. J. Teo, J. A. van Kan, and F. Watt: Appl. Phys. Lett. **88** 171106 (2006)
6. T. C. Sum, A. A. Bettiol, H. L. Seng, J. A. van Kan, and F. Watt: Appl. Phys. Lett. **85** 1398 (2004)
7. J. H. Jang, W. Zhao, J. W. Bae, D. Selvanathan, et al: Appl. Phys. Lett. **83** 4116 (2003)
8. M. L. von Bibra and A. Roberts: J. Light. Technol. **15** 1695 (1997)
9. D. M. Ruck, S. Brunner, K. Tinschert, and W. X. F. Frank: Nucl. Instrum. Methods B **106** 447 (1995)
10. K. Liu, E. Y. B. Pun, T. C. Sum, A. A. Bettiol, J. A. van Kan, and F. Watt: Appl. Phys. Lett. **84** 684 (2004)
11. P. D. Townsend, P. J. Chandler, and L. Zhang: Optical Effects of Ion Implantation, Cambridge University Press, Cambridge, 1994
12. M. L. von Bibra and A. Roberts: J. Canning: Opt. Lett. **26** 765 (2001)

Box 11: Tissue Engineering and Bioscience Methods Using Proton Beam Writing

J.A. van Kan

Tissue engineering is a rapidly developing and highly interdisciplinary field that applies the principles of cell biology, engineering, and materials science to the culture of biological tissue. The artificially grown tissue then can be implanted directly into the body, or it can form part of a device that replaces organ functionality.

In natural tissues, the cells are arranged with a three-dimensional (3D) organization which provides the appropriate functional, nutritional, and spatial environment. The culture of tissues 'in vitro' therefore requires biocompatible tissue scaffolds to give the cells not only physical support, but also to mimic the conditions of natural tissue, thereby enhancing cell proliferation and differentiation (the process by which cells reproduce to form a particular cell type from, e.g., stem cells) necessary for the regeneration and replacement of damaged cells. 3D amorphous mesh structures made from collagen or polymer fibres are common types of tissue scaffolds currently used [1]. These types of structures are aimed at providing the cells with the same organisation and mechanical properties of the natural tissue they replace. Although this approach has had some success in skin and bone production, mesh scaffolds do not exhibit the well-defined microscopic geometrical environment necessary for the strict geometrical guidance and confinement required by most cells.

Recently several groups have investigated the influence of different matrices and cell-surrounding media on tissue engineering; e.g., the use of platelet-rich plasma and mononuclear cells from bone marrow has shown potential for bone and dental implants [2]. A bovine collagen coated sponge was used to develop an acellular artificial dermis, composed of human collagen and glycosaminoglycan secreted by cultured human fibroblasts [3]. Kanai et al. [4] reported on the use of the human hepatocellular carcinoma cell line (FLC-4) cultured in a radial-flow bioreactor. From these examples it is clear that different cell lines require different conditions and

J.A. van Kan (✉)

Department of Physics, Centre for Ion Beam Applications, National University of Singapore, Singapore 117542, Singapore

e-mail: phyjavk@nus.edu.sg

extreme care has to be taken in preparing the most suitable environment for different cell lines.

Tissue scaffolds of specific 3D geometry may have the ability to promote and control, via topographic stimuli, tissue differentiation and creation of an extra cellular matrix (ECM), thereby increasing the quality and range of tissues suitable for graft. Current work has centred on the effects of surface topography on cell adhesion, motility, cell guidance, and cell shape characteristics; and there is a growing body of evidence to suggest that these characteristics can be influenced by surface topography [5–7]. Alterations to the morphological behaviour of cells have major functional consequences, in that changes in shape are closely associated with the cytoskeleton, which in turn is linked to transmembrane proteins (see chapter "Nanoscale Engineering in Biosciences"). Contact mediated processes regulate cell migration and proliferation, and these in turn are coupled to the mechanisms which control gene activity and therefore cell function. Cell function regulation through contact with specific surface topography has exciting possibilities.

Although it is well known that the behaviour and functions of cells are changed by geometric constraints and substrate surface properties, little work has been carried out on the effects of 3D microsubstrate geometry on cell behaviour. This knowledge is important for the success of tissue engineering, so that cells can organize in suitable 3D environments and function properly as an organ in vivo. The major reason for this lack of information stems from the general unavailability of precisely micropatterned 3D substrates. P-beam writing (see chapter "Proton Beam Writing: A New 3D Nano Lithographic Technique") is a technique which can produce 3D high aspect ratio micromachined surfaces of different shapes and patterns by altering a resist structure using a focused beam of MeV protons in a fast, direct write process. 3D microstructures with well defined geometrical shapes have been produced at sizes down to 100 nm, well below cell dimensions. By utilising a biocompatible resist material (e.g., polymethylmethacrylate [PMMA]), the technique of proton beam writing has extremely high potential for the production of substrates for cell and tissue engineering [8, 9].

A 3D PMMA resist structure was designed to test longitudinal movement along ridges: a circular ring structure containing a single 30-μm-wide, 20-μm-deep groove was interrupted with four radial 400-μm-long microridge outlets with widths of 5, 20, 25, and 30 μm, respectively. The micromachined PMMA substrate was sterilized following the protocol in [8]. (For other cell lines optimization of this procedure might be required.) In this structure, fibroblast cells migrated and spread outward until they reached the circular barrier. Cells were then observed to move along the three largest ridge channels (Fig. 1b–d), but not along the narrowest ridge of 5 μm (Fig. 1a). This result indicated that fibroblasts have restricted movement along a 5-μm-wide ridge, but can move easily along ridges of 20 μm or more. Interestingly, when cells pass along the ridges and move into the region outside the barrier, they then move back toward the main cell cluster, possibly sensing some chemical cues from the main body of cells still trapped within the circular barrier. Our results indicate that 3D microstructures can potentially deter fibroblasts from spreading and growing into a particular tissue environment. These 3D microstructures may be
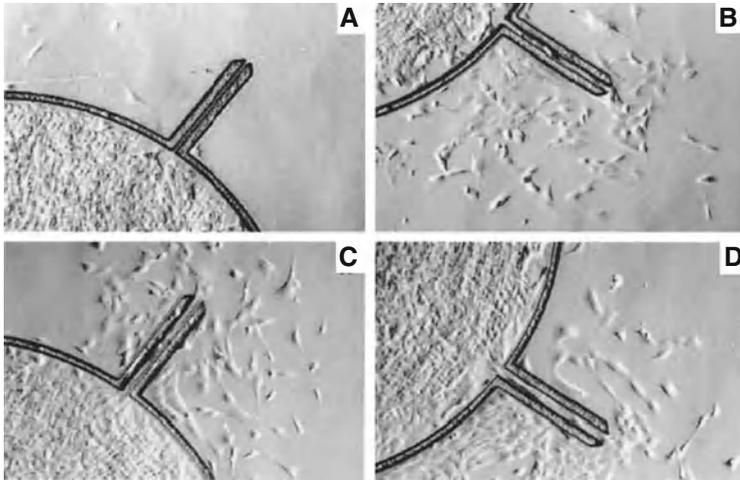
Fig. 1 Fibroblasts selectively migrate and spread through smooth surface channels. A modified 3D microstructure corral containing four smooth channels with widths of 5, 20, 25, and 30 μm. (**A–D**) Cell growth and spreading through the four channels on day 7. Channel width: (**A**) 5 μm; (**B**) 20 μm; (**C**) 25 μm; (**D**) 30 μm. The length of the channels is 400 μm

useful in tissue engineering when multiple cell types are needed to form a proper tissue, and fibroblast overgrowth can be detrimental in the proper formation and function of the tissue.

As a unique direct write technique, random and complex microstructures can be generated that may reflect more closely the in vivo tissue environment, facilitating the study of cell–cell and cell–matrix interactions for tissue engineering. In tissue engineering multiple substrate samples are often required. Since p-beam writing is a suitable technique to fabricate high quality metal masters (see Box 12) it is natural to use these masters to replicate multiple copies of the cell growth substrates in polymers, either through nanoimprinting or soft lithography [10].

References

1. R.P. Lanza, R. Langer, W.L. Chick, Principles of Tissue Engineering, Academic Press, New York, 1996.
2. C.H. Filho, I. Kerkis, A. Kerkis, et al. Artif. Organs 31 (2007) 268–273.
3. Y.K. Seo, K.Y. Song, Y.J. Kim, J.K. Park, Artif. Organs 31 (2007) 509–520.
4. H. Kanai, H. Marushima, N. Kimura, et al., Artif. Organs 31 (2007) 148–151.
5. P. Clark, P. Connolly, A.S.G. Curtis, J.A.T. Dow, C.D.W. Wilkinson, J. Cell Sci. 99 (1991) 73.
6. P. Clark, P. Connolly, A.S.G. Curtis, J.A.T. Dow, C.D.W. Wilkinson, Development 108 (1990) 635.
7. A. Curtis, C. Wilkinson, Biomaterials 18 (1997) 1573.
8. F. Sun, D. Casse, J.A. van kan, R. Ge, F. Watt, Tissue Eng., 10 (2004) 267.
9. J.L. Sanchez, G. Guy, J.A. van Kan, T. Osipowicz, F. Watt, Nucl. Instrum. Methods B 158 (1999) 185.
10. J.A. van Kan, L.P. Wang, P.G. Shao, A.A. Bettiol, F. Watt, Nucl. Instrum. Methods B 260 (2007) 353.

Box 12: Stamps for Nanoimprint Lithography

J.A. van Kan and K. Ansari

P-beam writing is a flexible direct write technique that is extremely suitable for rapid prototyping down to the nanometer level and therefore, in combination with electroplating, extremely suitable for the production of high quality masters for nanoimprinting. Direct write processes for a long time have been considered intrinsically slow and inefficient compared with masked processes for the mass production of large-area, high-density, low-dimensional structures. However, the increased technical complexities and predicted increased cost of producing feature sizes smaller than 100 nm has called into question the traditional role of masked processes as the universal method of mass production. Direct write processes therefore may have some distinct advantages when used in combination with nanoimprinting [1]. UV cured imprinting, employing a quartz stamp, has been successfully applied using step-and-flash imprint lithography. Here a photocurable prepolymer is cured and molded [2]. This technique has been mainly applied in semiconductor nanofabrication. Thermal nanoimprint lithography is a second form of nanoimprinting. Here either a Si or Ni mold is heated above the glass transition temperature of a polymer and pressed into the polymer to replicate features in the stamp [3]. Both these techniques have shown great success in low aspect ratio nanoimprinting. The p-beam written stamps have potential applications in all the developed technologies based on molds and stamps [1, 4, 5], such as template fabrication for single molecule electronics, nanophotonics, nanofluidics, biosensors, etc.

P-beam writing in combination with Ni electroplating has demonstrated the production of 3D Ni stamps [6]. Here we discuss the production of a high-quality, void-free, high aspect ratio metallic stamp of 100 nm width and $2\,\mu$m depth, using p-beam writing coupled with electroplating. P-beam writing exhibits the following features: (a) the ability to fabricate structures with smooth and vertical sidewalls, which is crucial to minimize pattern distortion during demolding in nanoimprint lithography; (b) the ability to fabricate high aspect ratio sub-100 nm features in a

J.A. van Kan (✉)

Department of Physics, Centre for Ion Beam Applications, National University of Singapore, Singapore 117542, Singapore

e-mail: phyjavk@nus.edu.sg

one-step process; and (c) since proton beam writing exhibits minimal proximity effects, it is ideally suited to produce features of high packing density. The fabrication of metallic stamps with these combined characteristics is very difficult to achieve with other lithography techniques such as focused ion beam (FIB), electron beam lithography (EBL), optical, or X-ray lithography.

Figure 1 shows a schematic representation of the process of stamp fabrication. This involves: (a) exposure of Si coated with PMMA, using p-beam writing; (b) deposition of a second metallization layer on the top surface which acts as a seed layer for the base of the stamp and development of the structures; (c) electroplating of the Ni stamp; and (d) nanoimprinting into a PMMA-coated Si wafer.

Fig. 1 Schematic representation of the process stamp fabrication (a–c) and nanoimprinting (d) using proton beam writing
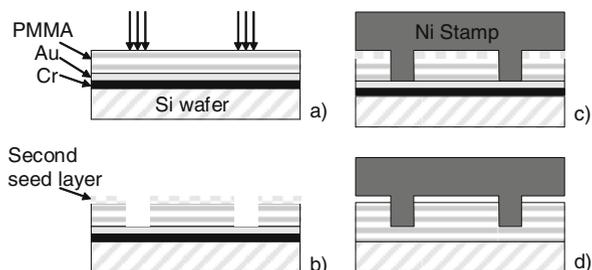
The fabrication process starts in a similar way as described in Sect. 3.3 of the chapter "Proton Beam Writing: A New 3D Nanolithographic Technique." The main difference is the fact that before resist coating a thin metallic seed layer for electrodeposition is applied (e.g., 20 nm Cr followed by 200 nm Au). In this example a layer of $2\,\mu m$ PMMA was spin coated onto a wafer. In the CIBA facility a beam of 2 MeV protons was focused to a spot size of $60 \times 90\,nm^2$ and magnetically scanned over an area of $50 \times 50\,\mu m^2$. The test pattern, a precursor to a microfluidic lab-on-a-chip system, consisted of 100 nm parallel trenches connected to reservoirs of $50\,\mu m$ (width) $\times 500\,\mu m$ (length).

After resist exposure and development the Ni electroplating was carried out using a typical Ni sulfamate bath solution with a sodium-dodecyl-ether-sulphate wetting agent, without organic additives, using a Technotrans AG, RD.50 plating system. The deposition is carried out employing first a low current density of $0.4\,A/dm^2$ for the first $50\,\mu m$, and then a high current density of $4\,A/dm^2$ for the next $200\,\mu m$. The initial low plating current density produces less intrinsic stress in the high aspect ratio structures, and is also coupled with the highest hardness. After delamination, the stamps were cleaned in toluene at $40\,^\circ C$ for 30 min. Figure 2a shows two electron microscope photographs of the test stamp featuring Ni ridges of 100 nm width and $2\,\mu m$ height.

To eliminate sidewall electroplating, thereby eliminating voids, two approaches have been explored: (1) deposition of the second metallization layer before proton beam patterning; this relies on the proton beam penetrating the second metallization layer without adverse effects; and (2) deposition of the second metallization

Fig. 2 (a) SEM images of a nickel stamp fabricated using proton beam writing and nickel electroplating. The stamp features several parallel ridges 100 nm wide and 2 μm depth. (b) Imprint of the Ni stamp in 8 μm thick PMMA spin coated on a silicon substrate, showing reproducible fine features, smooth sidewalls, and vertical structures
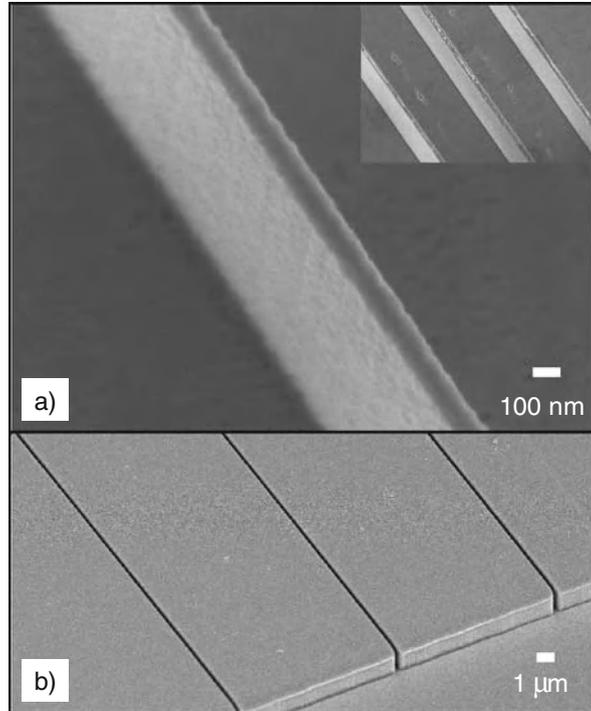
layer after proton exposure, but before developing the exposed areas. Both methods rely on the penetration of the developer through the second metallization layer, which has good adhesion to the resist. This layer should be thin enough to develop the structures but thick enough to provide a conductive layer for the subsequent over-plating process. Our experimental results show that deposition of a second metallization layer of 3 nm Ti provides adequate adhesion and good conductivity. Both processes outlined above were successful in complete elimination of voids. Hardness and Young's modulus of 5 and 213 GPa were obtained, respectively, which is consistent with the previously reported values [7]. A surface roughness of 7 nm RMS was measured with AFM.

The stamp with the 100 nm wall has been successfully used to replicate a negative of this structure in PMMA using a commercial nanoimprinter (Obducat Technologies AB, NIL-2-PL 2.5 in nanoimprinter [8]). The nanoimprinting was carried out using 8 μm thick PMMA resist 950 K molecular weight spincoated on a Si wafer. The imprinting was carried out using the masters directly without any treatment with an antiadhesion agent. A SEM photo of the imprinted pattern, which has vertical sidewalls and flat surfaces identical to the stamp, is shown in Fig. 2b.

References

1. S. Y. Chou., P. R. Krauss, and P. J. Renstrom, Appl. Phys. Lett. **67**, 3114 (1995).
2. M. Colburn, S. Johnson, M. Stewart, S. Damle, T. C. Bailey, B. Choi, M. Wedlake, T. Michaelson, S. V. Sreenivasan, J. Ekerdt, and C. G. Willson, Proc. SPIE-Int. Soc. Opt. Eng. **379**, 3676 (1999)
3. S. Y. Chou, P. R. Krauss, and P. J. Renstrom, Science **85**, 272, (1996).
4. N. Stutzmann, T. A. Tervoort, K. Bastiaansen, and P. Smith, Nature **407**, 613 (2000).
5. E. Schaffer, T. Thurn-Albrecht, T. P. Russell, and U. Steiner, Nature **403**, 24 (2000).
6. K. Ansari, J. A. van Kan, A. A. Bettiol, and F. Watt, Appl. Phys. Lett. **85**, 476 (2004).
7. T. Fritz, W. Mokwa, and U. Schnakenberg, Electrochimica Acta **47**, 55 (2001).
8. B. Heidari, I. Maximov, E.-L. Sarwe, and L. Montelius, J. Vac. Sci. Technol. **B 17** 2961 (1999).

Box 13: Silicon Micro/Nano-Fabrication Using Proton Beam Writing and Electrochemical Etching

Daniel J. Blackwood and Ee Jin Teo

The fabrication of 3D structures in silicon is at the heart of many state-of-the-art technologies. Conventional etching technologies require multiple processing steps to fabricate free-standing multilevel structures [1, 2]. The two main wet chemistry procedures are chemical etching in KOH, which is unfortunately isotropic with very slow etching of <111> planes, and electrochemical etching in HF [3]. The electrochemical etching rate is proportional to the flux of holes to the silicon/electrolyte interface, i.e., the applied current density, so it is easy to control and anisotropic. Etching Si in HF at low current densities produces porous silicon, which can be selectively removed by KOH. The low cost and easy implementation of electrochemical etching has made it an attractive alternative to dry etching. Selective patterning can be achieved through a mask or etch stop techniques [4]. Although undercutting of the mask is a problem, this can be improved in n-type silicon by prestructuring the masked surface in KOH to form a periodic array of etched pits [5, 6]. Unfortunately, there are disadvantages to this approach: the need for spatial periodicity; limitation to n-type silicon; and inability to fabricate multilevel structures.

Because the electrochemical etch rate depends on the flux of holes to the interface, the bias required for etching is proportional to the silicon's conductivity. Thus, controlling this laterally across the surface results in location dependant etch rates, allowing micromachining (Fig. 1). One way to locally control the conductivity of the silicon is to use focused ion beams, i.e., by ion implantation or creating damage. The former relies on selective doping with B, P, or Ga, followed by annealing to create n- and p-type regions [7]. The latter, creation of localized damage to reduce the carrier density, is achieved by irradiating with focused MeV protons or helium ions. These ions have a range of tens of microns in silicon and each ion creates displacement defects in the lattice along its trajectory. The defect generation rate is uniform until close to the end-of-range where there is a sharp increase. After

D.J. Blackwood (✉)
Department of Materials Science and Engineering, National University of Singapore, Singapore
e-mail: msedjb@nus.edu.sg
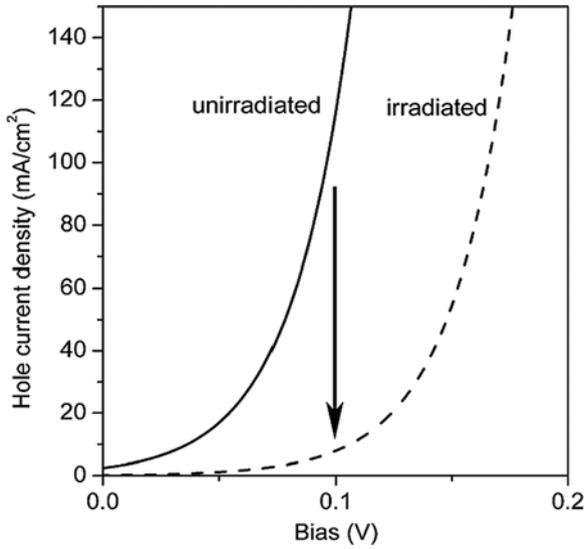
Fig. 1 I–V curve of low and medium resistivity p-type silicon undergoing anodization in HF. The reduction in current density at the irradiated region is indicated by the *arrow*
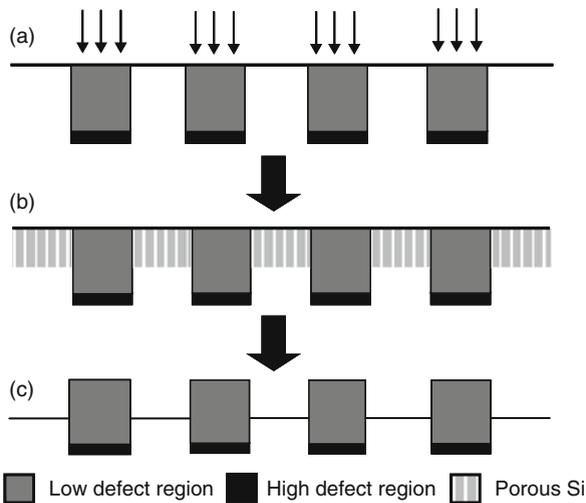
Fig. 2 Schematic of machining process: (**a**) Damage by focused ion beam irradiation (**b**) followed by electrochemical etching in HF (**c**) and removal of porous silicon using KOH. Figure 1 reproduced with permission from Teo et al. [8], Copyright Wiley-VCH Verlag GmbH & Co. KGaA

electrochemical etching the wafer surface is patterned with regions of porous silicon (Fig. 2).

MeV protons penetrate much deeper than the keV heavy ions used in previous irradiation studies that penetrate $\ll 1\,\mu$m. This deeper penetration enables the

fabrication of thicker patterned layers. Higher beam fluence at any region produces a higher damage concentration and thus a greater increase in local resistivity and a slower etching rate. Thus, pausing the beam at different locations allows any pattern of localized damage to be built up. The ability to control the depth of damage by varying the accelerating voltage introduces the possibility of producing 3D structures. Because MeV protons follow an almost straight path in silicon, it is possible to obtain structures with vertical sidewalls, overcoming the undercutting encountered with surface masks [5]. Finally, no post irradiation annealing is required because the process is based on damage rather than selective doping.

Figure 3 shows an SEM image of a uniform array of closely packed, high aspect ratio pillars obtained by single spot irradiations of a focused proton beam on a p-type wafer of resistivity 5 Ω cm. Each spot had an accumulated fluence of 5×10^{16} protons cm^{-2}. The sample was etched in 1:1:2 HF : H_2O : EtOH for 15 min at $40 \, mA \, cm^{-2}$. The pillars are $\sim 4.5 \, \mu m$ high with diameters of $0.6 \, \mu m$ and a periodicity of $2 \, \mu m$. The profile of the pillar reveals vertical and smooth sidewalls with slight broadening at the base, consistent with a straight path proton beam. Periodic arrays of submicron diameter pillars are potentially important for the fabrication of photonic crystals. Channeling the ion beam along a set of crystal planes reduces the probability of nuclear collisions with silicon atoms [9], leading to a significant reduction in damage close to the surface. This allows the fabrication of sharp needles that have tips with step slope angles of 85° and radii of curvature of $\sim 15 \, nm$ (Fig. 4). Such nanotips could find applications as scanning probe microscopes or for the mass transport of dots of molecules onto substrates.
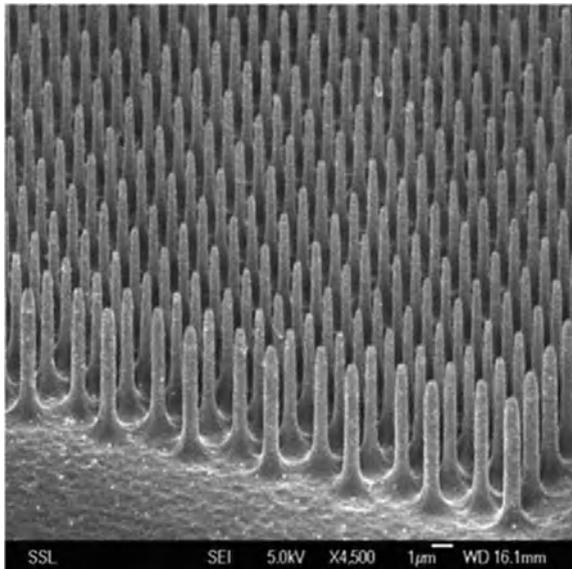
Fig. 3 Array of high aspect-ratio pillars obtained by single spot irradiations
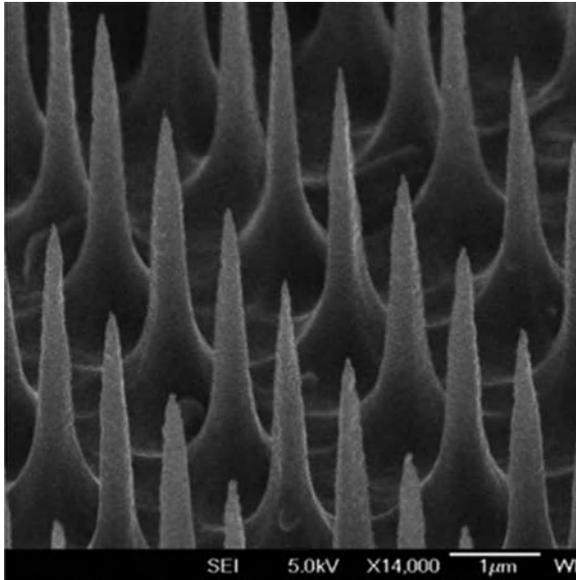
Fig. 4 Silicon needles obtained by channeling the beam along the $< 100 >$ crystal axis; radius of curvature of tip $\sim 15\,\text{nm}$

Figure 5 shows an array of tilted pillars, formed by irradiating with 1×10^{16} protons cm^{-12} at an angle of $15°$ off-normal and etching at $30\,\text{mA cm}^{-2}$ for 15 min. The pillars correspond to the paths taken by the protons through the silicon, which leads to the conclusion that the etching profile is dominated by the low defect region rather than the deeper end-of-range high defect region. Recent modeling studies confirmed this deduction. It would be impossible to produce such tilted structures using mask-based etching techniques in a single step.

Since protons only penetrate the silicon lattice to a finite depth, it is possible to etch beyond this range to expose the undamaged lattice. This means that can-tilevers and multilevel structures can be created in a single etch step by exposing the silicon with two different proton energies. The regions irradiated with lower energy, penetrated to a shorter depth, will begin to undercut first whilst the areas with higher energy irradiation continue to increase in height. To demonstrate this capability, a bridge structure was irradiated using 0.5 MeV protons for the beam and 2 MeV protons for the supporting pillars. Figure 6a schematically shows the subsequent etching process to form a free-standing bridge. Initially, etching occurs only in the unirradiated regions, but when the etching proceeds beyond the end-of-range of 0.5 MeV protons ($\sim 6\,\mu\text{m}$), undercutting of the beam starts whilst the pillars continue to grow. Etching to a depth of $14\,\mu\text{m}$, the bridge's beam is fully undercut and separated from the substrate, being supported by the pillars (Fig. 6b). Further etching to $25\,\mu\text{m}$ revealed the complete structure of a free-standing bridge; undercutting of the pillars did not occur because the range of 2 MeV protons is $48\,\mu\text{m}$ (Fig. 6c).
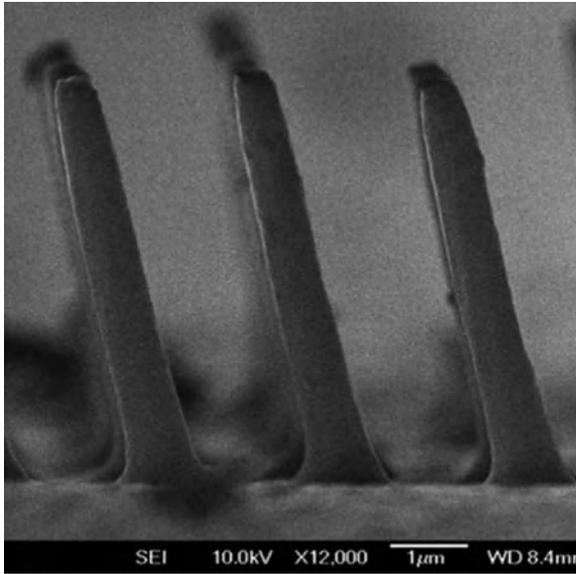
Fig. 5 Tilted pillars formed by tilting the wafer at a 15° angle with respect to the incident proton beam. Figures 3 and 4 reused with permission from Teo et al. [10]. Copyright 2004, American Institute of Physics. Figure 5 reprinted with permission from Breese et al. [11]. Copyright (2006) by the American Physical Society
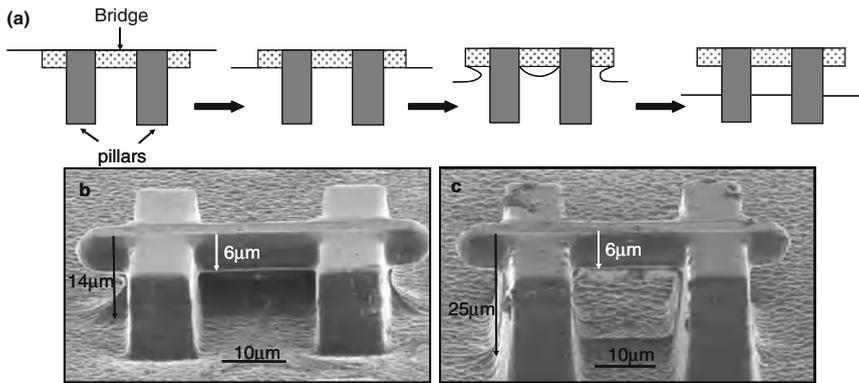
Fig. 6 (**a**) Evolution of the double-energy irradiated structure with etching depth. (**b**) At etch depth of $14\,\mu$m, the bridge starts to separate from the substrate. (**c**) The bridge is completely free-standing at an etch depth of $25\,\mu$m. Reused with permission from Teo et al. [10]. Copyright 2004, American Institute of Physics

References

1. H. G. Craighead, Science **290**, 1532 (2000).
2. A. N. Cleland and M. L. Roukes, Appl. Phys. Lett. **69**, 2653 (1996).
3. V. Lehmann and U. Gosele, Appl. Phys. Lett. **58**, 856 (1991)

4. A. G. Nassiopoulos, Properties of porous silicon, edited by L. Canham, No. 18, (INSPEC, London, 1997) pp. 77–80.
5. V. Lehmann and H. Foll, J. Electrochem. Soc. **137**, 653 (1990).
6. P. Kleimann, J. Linnros and R. Juhasz, Appl. Phys. Lett. **79**, 1727 (2001).
7. J. Xu and A. J. Steckl, Appl. Phys. Lett. **65**, 2081 (1994).
8. E. J. Teo, M. B. H. Breese, A. A. Bettiol, D. Mangaiyarkarasi, F. Champeaux, F. Watt and D. J. Blackwood. Multicolor photoluminescence from porous silicon using focused, high-energy helium ions, Adv. Mater. **18**, 51–55 (2006).
9. L. C. Feldman, J. W. Mayer and S. T. Picraux, Materials Analysis by Ion Channeling, (Academic Press, New York, 1982).
10. E. J. Teo, M. B. H. Breese, E. P. Tavernier, A. A. Bettiol, F. Watt, M. H. Liu and D. J. Blackwood, Appl. Phys. Lett. **84**, 3202 (2004)
11. M. B. H. Breese, F. J. T. Champeaux, E. J. Teo, A. A. Bettiol and D. J. Blackwood, Phys. Rev. B **73**, 035428 (2006)

Nanoscale Materials Modification for Device Applications

Robert G. Elliman

This chapter considers device applications of ion beams that involve nanoscale modification of materials. It is necessarily selective and does not include broad classes of applications such as ion beam assisted deposition of thin films or tribological modification of surfaces for improved wear, corrosion resistance or biocompatibility. Instead it aims to illustrate the diversity of ion beam applications by providing some contemporary examples that exploit modification of electronic, optical, magnetic, and mechanical properties of materials.

No attempt is made to provide an historical overview of various applications, nor to provide a comprehensive literature survey. Instead, the reader is referred to the broader literature for such information.

1 Microelectronics

Ion beam techniques, such as ion implantation, reactive ion etching, and plasma enhanced deposition, have been instrumental in the microelectronics revolution and have continuously been adapted to the challenges presented by device scaling, the introduction of new materials (e.g., SiC, ZnO, GaN, diamond, etc.), and the development of new device structures.

1.1 Ion Implantation

The single most significant application of ion beams in microelectronics is ion implantation. To put this in context, a large-scale device manufacturer would aim to process up to 50,000 wafers/month, with each wafer requiring 20–30 implants. This typically involves the use of around 20 implanters, each with the capacity to implant more than 250 wafers/h. As a result, there are currently around 4000 ion implanters employed worldwide for routine semiconductor processing.

R.G. Elliman (✉)
Electronic Materials Engineering Department, Research School of Physics and Engineering, The Australian National University, Canberra ACT 0200, Australia
e-mail: rob.elliman@anu.edu.au

The most common application is the manufacture of silicon-based devices where implantation is used primarily for doping. This application employs ion energies in the range from 200 eV to 3 MeV, and fluences in the range from 10^{11} cm^{-2} for threshold voltage adjustment to several 10^{16} cm^{-2} for contact and well formation, as shown in Fig. 1. Other significant applications include doping and
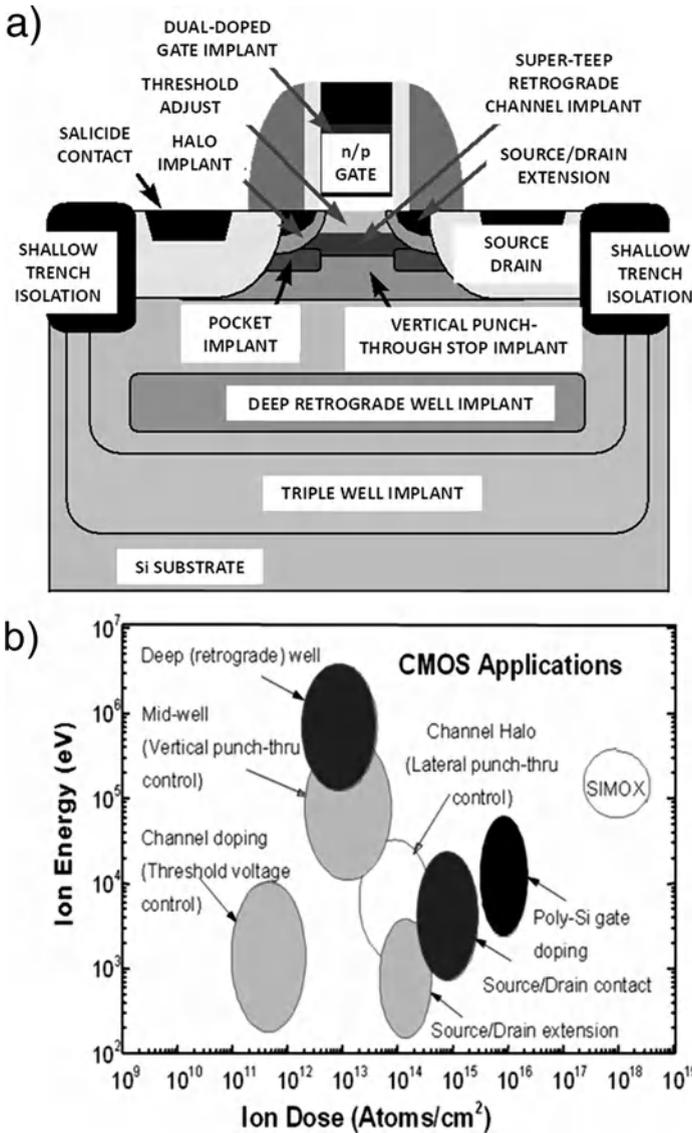
Fig. 1 (**a**) Schematic of a modern CMOS FET indicating the various implants involved in its manufacture, (**b**) range of implant energies and fluencies employed for modern device manufacture. (After J. Poate, private communication)

electrical isolation of III-V semiconductor devices, and custom applications such as Smartcut® and SIMOX® for producing silicon-on-insulator material.

Ironically, as semiconductor devices have shrunk, the energy range of implanted ions employed for their manufacture has increased. At the lower end, energies of around 200 eV are employed to dope near surface structures such as the source-drain extension and contact regions of transistors, whilst at the upper end, MeV ions are employed to dope layers that extend beneath the device, such as deep wells. At the low energy end of this spectrum the implanted ions have ranges of a few nanometers and accurate control of dopant distributions during implantation and subsequent processing is a critical challenge [1].

1.2 Single Ion Implantation

1.2.1 Doping Statistics

As electronic devices continue to shrink, the area of active elements becomes so small that they contain countably few dopant atoms. The statistical variation between devices then becomes a major limitation to device performance. This problem was recognized as early as 1975 by Keyes [2], who noted that statistical variations in the number and placement of dopants could be a major source of device variability, even in devices much larger than those now in production.

It has only recently become possible to test these concepts experimentally [3, 4], with the first experiments being performed by Shinaka et al. [3] using a single-ion implantation process to precisely position ions within the channel regions of field effect transistors. Figure 2 compares the threshold voltage distributions of devices fabricated with ordered and disordered dopant distributions. Devices with ordered dopant were found to have reduced threshold voltage variation, as predicted [2]. This
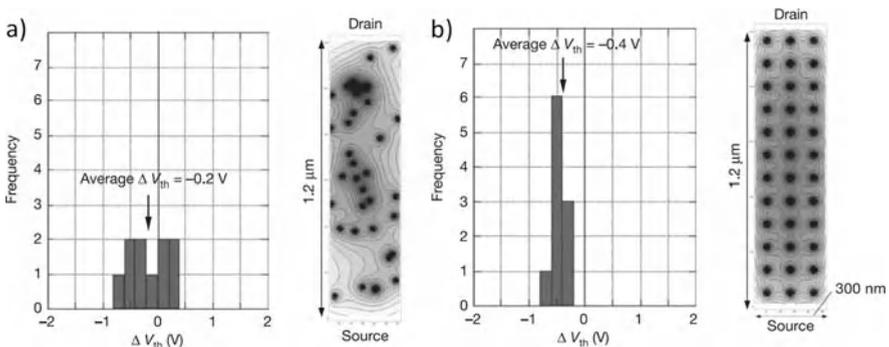
Fig. 2 Threshold-voltage histograms and dopant distribution maps for field effect transistors with gate regions doped with (**a**) a random distribution of dopants, and (**b**) an ordered array of dopants. The standard deviation of the latter is 0.1 V, three times smaller than for the devices doped with random distrubutions (From [3])

single-ion implantation technique has great potential for improved device performance, but will require significant development to become a large scale production technique.

1.2.2 Quantum Computing

The ability to implant single ions at precise locations also forms the basis of attempts to develop a solid-state quantum computer based on the interaction between neighboring P ions in silicon [5]. Proposed schemes for P-ion placement include implantation through apertures defined by conventional lithography or scanning-probe positioning of nanoapertures [6], as illustrated in Fig. 3. Detection of the individual ions can be achieved by employing the substrate as a charge detector [7] or by detecting secondary electron emission from the implanted ion [3].

Whilst a quantum computer based on implanted dopants has not yet been demonstrated, quantum interactions such as a Coulomb blockade have been reported for charge transport through single-electron transistor structures containing P-doped "islands" [8], as shown in Fig. 4. In this case the Coulomb blockade effect results from charging of the doped island. However, attempts are being made to extend such measurements to the few-electron regime with the aim of observing and manipulating discrete quantum states.

Fig. 3 Schematic of a scanning probe aperture employed for single-ion implantation showing: (1) ion beam; (2) electrostatic lens element; (3) pre-collimating aperture; and (4) a cantilever with the final aperture (From [6])
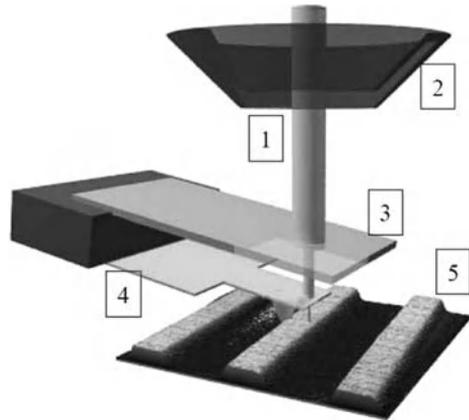
1.3 Bandgap Engineering

1.3.1 Quantum Well Intermixing

Many modern electronic and optoelectronic devices exploit engineered electronic structures produced by combining materials with different bandgaps. For example, the heterostructure shown in Fig. 5 produces a quantum well defined by the bandgap differences between GaAs and AlGaAs. The energy levels responsible for electronic
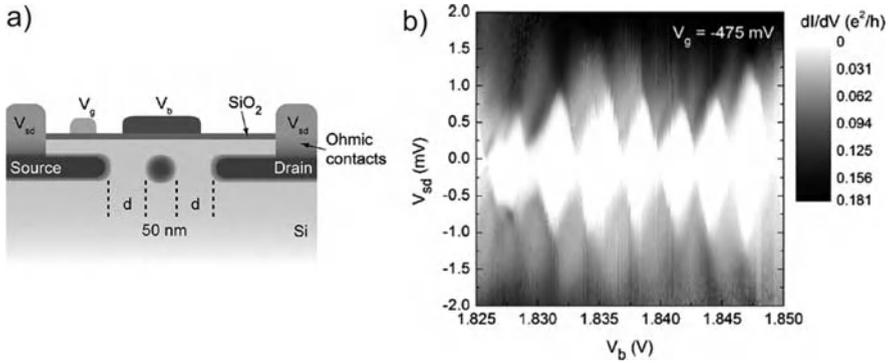
Fig. 4 (**a**) Schematic of a test device in which a doped island containing approximately 600 implanted phosphorus ions is positioned between two electrodes. (**b**) Low temperature transport properties showing Coulomb diamonds for a device with 35 ± 5 nm gaps between the island and source and drain contacts (From [8])

and optical transitions can be estimated from a "particle-in-a-box" model using the known potential distribution. However, these levels are very sensitive to the shape of the potential well. As a consequence, small changes in shape can be used to tune the levels and thereby control device parameters such as the emission wavelength of a laser [9]. This can be achieved most simply by interdiffusing or intermixing neighboring layers, as shown in Fig. 5.

Intermixing can be achieved by conventional diffusion, but this generally requires high processing temperatures and affects all devices on a wafer simultaneously. In contrast, ion irradiation offers a means of inducing local intermixing. This relies on radiation- or impurity-enhanced diffusion of layers and therefore depends on the specific material system and irradiation and annealing conditions employed.

Ion beam induced intermixing has been applied to a wide range of devices based on quantum wells, quantum wires, and quantum dots [10]. Specific examples will be discussed further in Sects. 2.2 and 2.3.
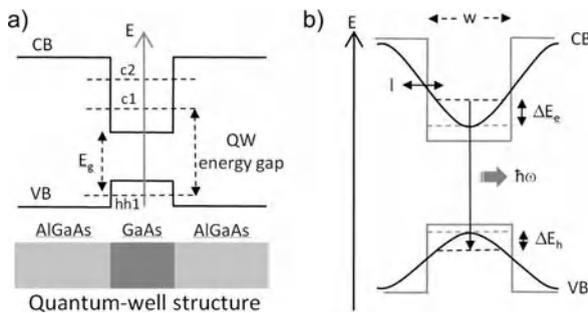
Fig. 5 (**a**) Schematic of an AlGaAs/GaAs/AlGaAs quantum well and the associated square well produced by the different bandgaps of GaAs and AlGaAs. (**b**) Schematic of the potential well before (*dotted line*) and after (*solid line*) intermixing

1.3.2 Alloys and Compounds

The electronic structure of semiconductors can be directly controlled by the implantation of high concentrations of impurities, either through alloying, as in the case of Ge implanted Si [11] or Mn doped Ge [12], or via the formation of compounds such as metal silicides [13, 14] or dielectrics, such as silicon dioxide [1]. The latter is a commercial process (SIMOX®—Separation by IMplantation of OXygen) used for the production of silicon-on-insulator wafers).

As one example of this approach, Fig. 6 depicts a field effect transistor with a GeSi strained-layer channel fabricated by Ge implantation and solid phase epitaxy [11]. This particular device showed improvements of up to 20% in channel hole mobility and transconductance when compared to a Si-only device, as shown in Fig. 6. This translates to a similar increase in the operating speed of the device and clearly demonstrates the potential of ion beam synthesis for improved device performance.
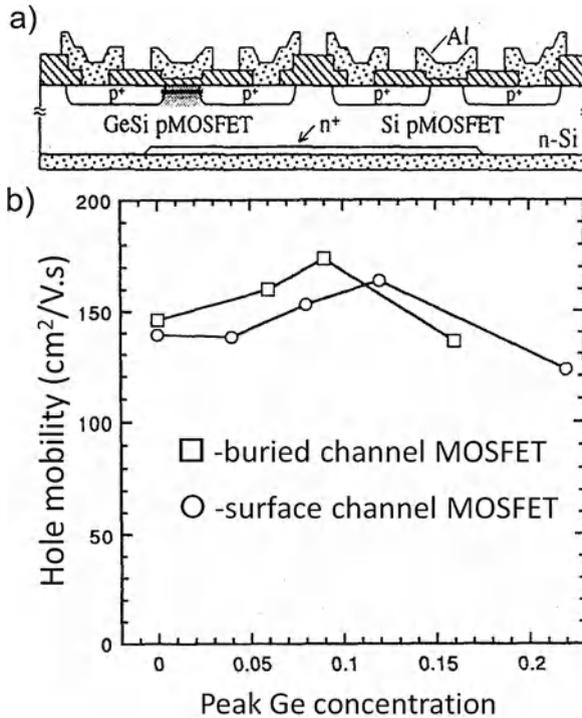
Fig. 6 (**a**) Schematic of a metal-oxide-semiconductor field effect transistor (MOSFET) fabricated from a GeSi strained layer produced by Ge-implantation of Si, and (**b**) the corresponding hole mobility measured for buried- and surface-channel devices (After [11])

1.4 Nanocrystal Memory

An important class of nonvolatile flash memory is based on the use of floating gate transistors, in which the logic state is determined by the threshold voltage
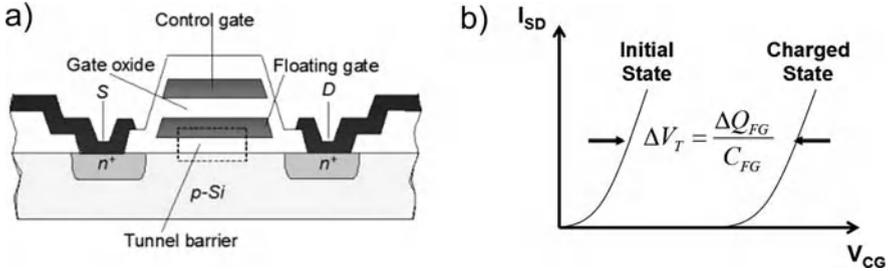
Fig. 7 (**a**) Schematic of a floating gate transistor. (**b**) Schematic diagram showing the shift in the current–voltage characteristics caused by charge on the floating gate. Symbols: S = Source; ΔV_T = change in threshold voltage; ΔQ_{FG} = change in charge on the floating gate; C_{FG} = capacitance of the floating gate

for transistor action, as shown in Fig. 7 [15, 16]. The threshold voltage is changed by introducing or removing charge on the floating gate. Early devices employed polycrystalline-silicon for the floating gate, but these were susceptible to memory loss due to charge leakage from the floating gate via defects in the tunnel barrier or gate dielectric [16]. Subsequent designs overcome this problem by using defects or nanocrystals as discrete charge traps [16–18]. This prevents complete charge loss in the event of dielectric failure and makes for more robust memory devices.

The first nanocrystal-based memory was demonstrated by Hori et al. [17] and employed Si nanocrystals (NCs) produced in a thin SiO_2 gate dielectric by ion beam synthesis. This device showed significant improvement over conventional memories, including a relatively low operation voltage, high endurance, fast write/erase speeds, and better charge retention. Although alternative means of nanocrystal synthesis were subsequently employed for commercial production, ion beam synthesis remains an active area of research for prototyping devices based on new and emerging materials, such as high-k dielectrics.

1.5 Impurity Gettering

Gettering is a process by which deleterious impurities are removed from the active regions of a semiconductor device by trapping them elsewhere in the material [1]. Trapping of impurities in Si is usually achieved by strain created by impurities or defects, or by second phases such as SiO_2 or amorphous silicon. The gettering process is broadly classified as intrinsic or extrinsic depending on whether it occurs throughout the bulk wafer (intrinsic) or in a locally modified region (extrinsic). Ion implantation has been used for extrinsic gettering since the very early days of the semiconductor industry, exploiting both implanted impurities and radiation damage to trap impurities. It underwent renewed interest when it was demonstrated that nanoscale voids produced by He irradiation of Si could act as very effective gettering sites [19].

The effectiveness of voids as gettering sites for metallic impurities is illustrated in Fig. 8 [20]. A silicon wafer was implanted with 100 keV H^- and 70 keV Cu^- ions.
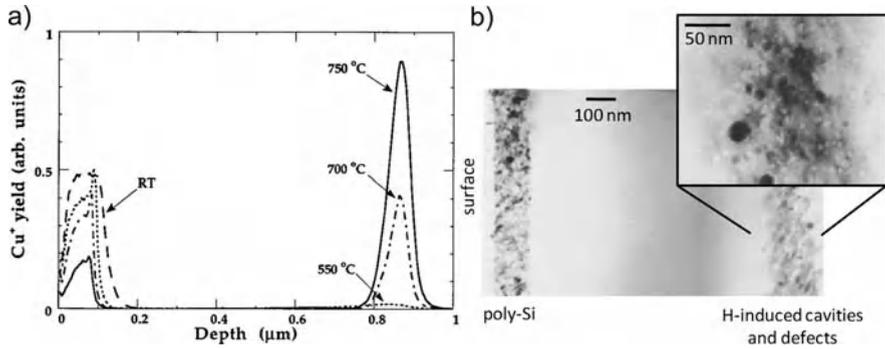
Fig. 8 (**a**) Cu distribution measured by secondary-ion mass spectrometry (SIMS) as a function of annealing temperature for a Si wafer implanted with Cu (70 keV, 3×10^{15} cm^{-2}) and H (100 keV, 3.2×10^{16} cm^{-2}). (**b**) XTEM image of the sample after H and Cu implants and annealing at 750 °C for 30 min (After [20])

As the sample is annealed to higher temperatures, the Cu is observed to migrate from the near surface into the Si wafer, and to trap at the voids produced by a deeper H implant. This has proven to be an effective means of gettering impurities from the near surface region of wafers.

2 Photonics

Ion beams are used extensively in the development, prototyping, and fabrication of optical devices and structures, both to modify the intrinsic properties of materials (e.g., refractive index, absorption, electronic structure) and to fabricate material structures (e.g., waveguides, photonic crystals, micro-electro-mechanical systems (MEMS)).

2.1 Waveguides

2.1.1 Refractive Index Modification

Refractive index changes result from radiation damage or compositional changes produced by ion implantation, and as a consequence ion implantation can be used to fabricate both passive and active waveguides in a wide range of linear and nonlinear optical materials, with the first studies dating back to the late 1960s. Applications include simple waveguides for the transport, splitting, and coupling of light as well as optically active guides used for second harmonic generation (SHG), optical parametric oscillators (OPO), amplifiers (OPA), and lasers [21]. Implantation can also be used to improve the photorefractive response of materials, thereby enabling more effective laser writing of photonic structures.

A significant advantage of ion implantation is the fact that refractive index changes can be defined by an implantation mask without physical etching of

the materials. This has particular advantages for high-index materials where edge roughness caused by etching is a major source of optical losses through scattering. The ability to modify the local refractive index also allows more complex planar structures to be fabricated, including gratings, mirrors, and optical cavities.

As an example, Fig. 9 shows a planar waveguide fabricated in $LiNbO_3$. The waveguide was defined by ion irradiation with 1.4 and 1.5 MeV He ions, with the dual implant used to produce a sufficiently thick barrier layer between the guide and substrate. A photonic bandgap (PBG) structure was subsequently fabricated within the waveguide by electron beam lithography and chemical etching. Characterization of the waveguide before the addition of the PBG structure shows several well-defined propagation modes, confirming the efficacy of the waveguide [22].
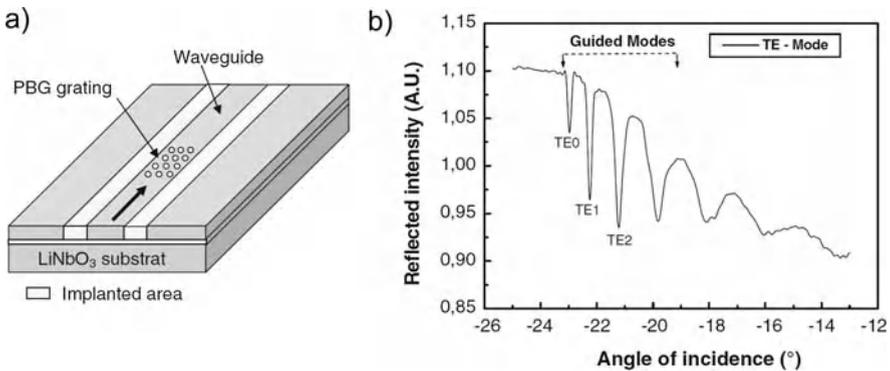
Fig. 9 (**a**) Schematic diagram of a lithium niobate waveguide containing a photonic crystals structure; (**b**) reflectivity of ion implanted lithium niobate as a function of angle shows transverse electric (TE) guided modes (From [22])

2.1.2 Bandgap Control for Low Loss Guiding

The operation of semiconductor lasers is often limited by absorption at the mirror facets that define the laser cavity. Excess absorption can cause catastrophic mirror damage and result in complete device failure. This can be overcome by employing quantum-well intermixing (QWI) to create active and passive sections in the same laser cavity, as shown in Fig. 10. Here, the QW of a GaAs/AlGaAs laser is intermixed to increase the bandgap of the material in the vicinity of the mirrors. The reduced optical absorption in this region allows the laser to operate at much higher power than conventional devices.

2.2 Switching and Detection of Light

Ion irradiation modifies the optical absorption of materials by introducing defects and impurities with specific absorption characteristics, changing the composition
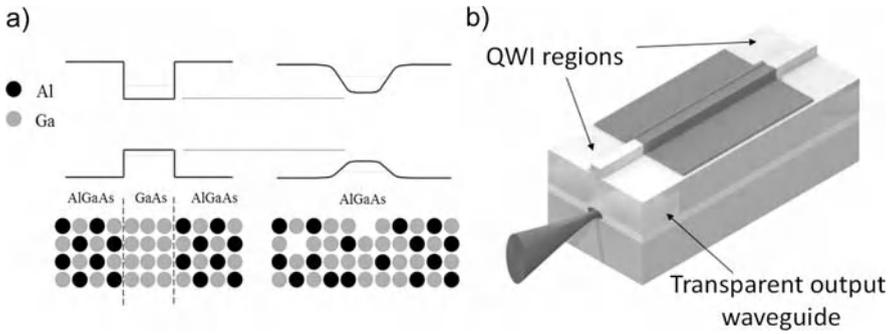
Fig. 10 (**a**) Schematic of the quantum well intermixing (QWI) process. (**b**) Schematic of a heterojunction laser showing the region in which intermixing has been applied (From www.intenseco.com/technology/)

of the material, forming new phases or inclusions, changing the carrier dynamics of the material, etc. Examples include the synthesis of metallic and semiconducting nanocrystals in glasses to produce characteristic linear and nonlinear absorption; quantum well intermixing of semiconductor heterostructures to tune the absorption characteristics of quantum well, wire, and dot photodetectors; and defect production in semiconductor-based saturable absorbers to improve their temporal response.

2.2.1 Metallic and Semiconducting Nanocrystals

The linear optical response of glasses containing metallic and semiconducting nanoparticles has been studied extensively and represents one of the earliest optical applications of ion implantation. The materials are formed by implanting glasses such as silica with appropriate elements using high-fluence ion implantation. During implantation or subsequent annealing the impurities precipitate to form nanometer-sized crystallites, as shown in Fig. 11. In the case of metallic particles, the glasses display a characteristic color due to resonant absorption which arises from a surface plasmon resonance exhibited by small metal particles [23, 24]. Such a resonance is evident in the optical absorption spectra shown in Fig. 11, and this can be tuned by controlling the composition and size of the nanocrystals. Similarly, semiconductor nanocrystals exhibit size-dependent absorption, which in this case is due to changes in their bandgap as a result of quantum confinement effects, a process that also gives rise to tunable luminescence from these materials.

Glasses containing metallic and semiconducting nanocrystals have been shown to exhibit interesting nonlinear optical behavior, including a third-order (Kerr) nonlinearity that gives rise to an intensity dependent refractive index [24, 25]. This response arises from size-dependent changes in the electronic structure of nanoparticles, carrier depletion and band filling, temperature rises, and phase changes. Such responses are of particular interest for making all-optical switches and modulators.
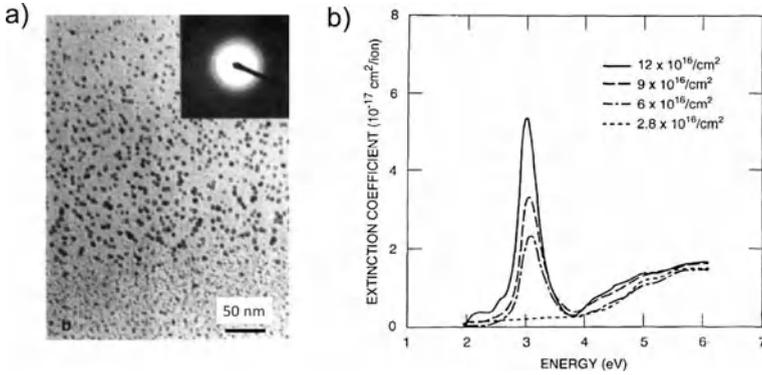
Fig. 11 (**a**) Transmission electron microscope image of Ag precipitates formed in silica by irradiation with 1.8 MeV Ag^+ (6×10^{16} cm^{-2}). (**b**) Optical absorption spectra of silica samples implanted to different Ag fluences (From [26])

2.2.2 Semiconductor Photodetectors

As discussed in Sect. 2.3, ion irradiation of semiconductor heterostructures (quantum wells, wires, and dots) can cause intermixing of neighboring layers and thereby change the electronic energy levels within the well. As a result, ion beam induced intermixing can be used to tune the transition energies for light absorption and emission. Moreover, because intermixing can be applied locally, it provides the basis for making devices with multiwavelength response, i.e., a single device can be modified to provide regions of different response, or individual devices on a single chip can be modified to have different responses.

As an example, Fig. 12 shows how the response of a quantum well infrared photodetector (QWIP) can be modified by appropriate intermixing [27]. The active region of the detector consisted of 50 AlGaAs/GaAs/AlGaAs quantum wells located
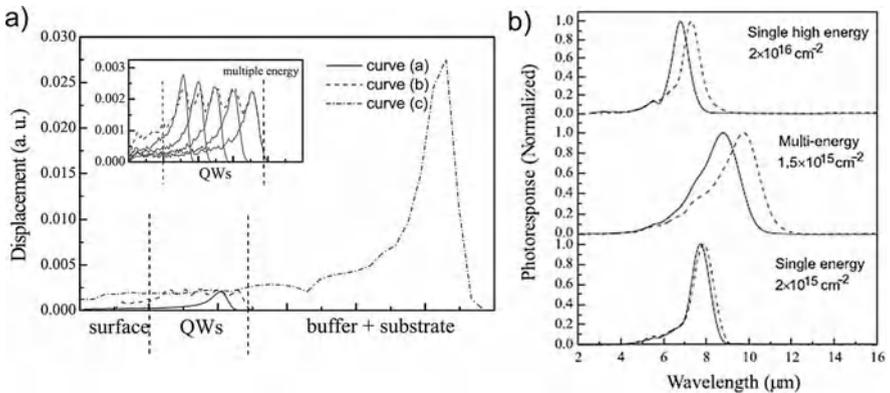
Fig. 12 (**a**) Displacement profiles produced in a quantum well infrared photodetector by different implantation schemes, (**b**) resulting photoresponse of the detector (From [27])

beneath a capping layer. Several different implant schemes were employed for inter-mixing in an attempt to maximize the response of the detector to long wavelengths. Specifically, the photoresponse was compared for three different implant schemes: a) a single irradiation in which the peak of the damage profile was located within the QW structure; b) a multiple-energy implant scheme designed to produce a uniform damage distribution throughout the QW region; and c) a single high-energy irradiation in which the peak of the damage distribution is well beyond the QW structure. These schemes are compared in Fig. 12, which also shows the photoresponse of the detector. In this case, the greatest shift in response was observed for the multiple energy implants.

2.2.3 Saturable Absorbers

Saturable absorbers are materials whose absorption coefficient decreases at high light intensity, a process sometimes referred to as photo-bleaching. These materials have a wide range of applications, including their use in semiconductor saturable-absorber mirrors (SESAMs), structures that are used to produce laser pulses for passive mode-locking and Q-switching of lasers [28, 29]. In this application, a thin (~ 10 nm) saturable absorber layer is incorporated within a Bragg mirror, as shown in Fig. 13. The absorber reduces the reflectivity of the mirror for low pulse intensities but increases the reflectivity for fast, high intensity pulses.

The saturation in absorption arises from depletion of electrons in the valence band and accumulation of electrons in the conduction band (Pauli blocking). Recovery of the initial absorption level is then controlled by the various carrier thermalization and relaxation processes, as shown in Fig. 13.

In this application, ion implantation is used to create defects that act as efficient recombination centers, thereby reducing the carrier lifetime and shortening the response time of the absorber. Using this approach, the response time of GaAs
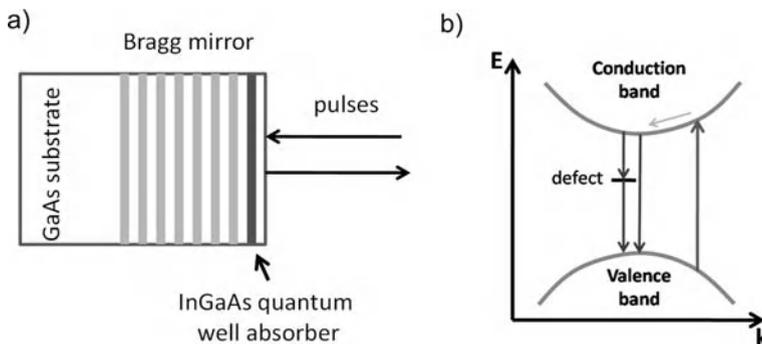
Fig. 13 (**a**) Schematic of a GaAs/AlGaAs Bragg mirror containing an InGaAs saturable absorber layer. (**b**) Energy band diagram for a direct-gap semiconductor showing excitation and relaxation mechanisms. Ion implantation is used to increase the rate of interband relaxation due to the introduction of defect-related mid-gap states

and InGaAs devices has been reduced by more than an order of magnitude, from several 100 ps to around 10 ps [30, 31].

2.3 Light Emission and Detection

Light emitting devices play a key role in modern photonics, and there is enormous research effort devoted to the development of both discrete and integrated devices. Ion implantation has played an important role in many of these studies, both as a prototyping tool and as a potential production tool. Major applications include: a) the introduction of optically active impurities such as Er into photonic structures, b) quantum well intermixing for tuning the emission or absorption wavelength of lasers, and c) the synthesis of new materials that produce or enhance light emission.

2.3.1 Defects and Doping

Ion irradiation of semiconductors and dielectrics produces a rich tapestry of defects, many of which are of interest as light emitting centers. For example, in Si both the W-centre and dislocation loops produced by B implantation of silicon [32] are of interest as potential light sources in silicon-based photonics, whilst the NV (nitrogen vacancy) centre in diamond, produced by nitrogen irradiation of diamond or vacancy production in N-containing material, is of interest as a single photon source and as a possible qubit for quantum computing [33].

The implantation of optically active impurities can also be used to control the optical emission from materials. An example of particular interest to the telecommunications industry is Er-doped materials capable of emitting light at $1.5\,\mu$m. This wavelength corresponds to the minimum propagation loss in silica fibers and is used for most long-haul, ground-based telecommunications. Amplification of signals is achieved by all-optical Er-doped fiber-amplifiers consisting of around 10 m of Er-doped silica fiber. The interest in Er-doped devices is largely driven by the prospect of developing more efficient sources and more compact amplifiers.

As an example of such applications, Fig. 14 shows a recent laser design based on a microdisk resonator that has been ion-implanted with Er [34]. This device was fabricated from an oxidized Si wafer by implanting the $1\,\mu$m thick oxide layer with 2 MeV Er ions, defining the microdisks using conventional lithographic and chemical etching techniques, and then undercutting the disks by selective reactive ion etching of the Si using XeF_2 gas. The device was used to demonstrate Er-based microdisk lasing and to study fundamental aspects of Er emission in these microcavities.

2.3.2 Semiconductor Laser Tuning

As discussed in Sect. 1.3, quantum well intermixing has been applied to a broad range of devices. For optoelectronic devices the main purposes of intermixing are to: a) confine light emission or lasing to particular regions of the device, b) concen-
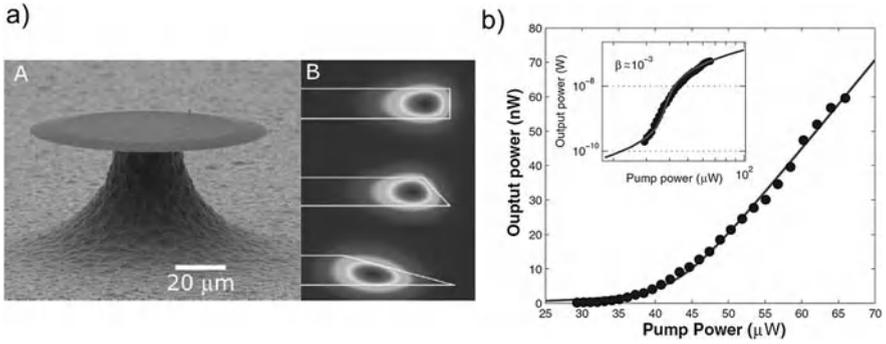
Fig. 14 (**a**) SEM image of an Er-implanted SiO$_2$ microdisk sitting on a Si pillar, and finite element simulation of the light intensity profile in the disk for different edge profiles. (**b**) The output power as a function of pump power, showing a lasing threshold at a pump power of 43 μW (From [34])

trate current injection in active regions of the device, and c) adjust the emission or absorption wavelength of the device.

As an example of how all three processes might be used, Fig. 15 shows a schematic diagram of a quantum wire laser formed by epitaxial growth of active layers on a V-groove–etched GaAs substrate. The quantum wire is formed from the slightly thicker region of the quantum well that forms at the bottom of the V-groove. Emission from the as-grown structure occurs from the entire QW layer, even though the emission wavelength is slightly different from the thicker wire. The V-shaped emission is clearly evident in the edge emission. By tilting the wafer during implantation it is possible to selectively intermix the active quantum well layers to eliminate lasing from regions other than the wire, as illustrated schematically in Fig. 15. Emission from the structure after such processing shows discrete emission from individual wires. A subsequent isolation implant can then be used to increase
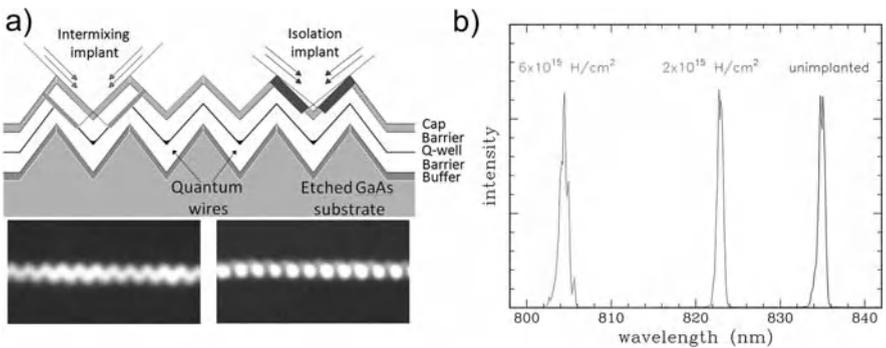
Fig. 15 (**a**) Schematic of a quantum wire laser array, and light emission intensity from the edge of the structure before intermixing, and after intermixing. (**b**) Emission from an intermixed GaAs/AlGaAs quantum well (From H. Tan, private communication)

the resistivity of the surface capping layer in order to concentrate current injection in the vicinity of the active wire.

Finally, it is possible to irradiate the active quantum wells, wires, or dots to induce intermixing and thereby modify the emission wavelength of the laser. Figure 15 also shows the measured emission wavelength from a quantum well structure after irradiation with protons, highlighting the shift to shorter wavelength as the effective width of the potential well is reduced.

2.3.3 Light Emitting Composites

The ability to synthesize new phases and inclusions in a wide variety of host materials provides a basis for making new light emitting composite materials. In recent years particular effort has been devoted to silicon-based light emission, where strategies include doping silicon with optically active impurities, such as rare-earth ions (e.g., Er) [35]; the synthesis of new phases such as silicides (e.g. Fe_2Si_3) [36]; direct modification of the band structure of silicon by alloying (e.g., alloys of Si with Ge, Sn and/or C) [37]; and the use of defects, such as point-defect clusters and dislocations [32]. Such approaches remain active areas of research and prototypic devices have been demonstrated in some cases.

A major change in the emphasis of this research was brought about by the discovery of strong visible luminescence from porous silicon [38], a sponge-like form of silicon produced by electrochemical etching, and from silicon nanocrystals produced by ion implantation [39]. This led to an intensive investigation of the properties and applications of silicon nanocrystals and culminated in a report of optical gain in a silica waveguide containing silicon nanocrystals [40]. During this period, silicon nanocrystals were also found to act as effective sensitizers for the excitation of optically active dopants such as Er, further stimulating the interest in such materials. This work has been the subject of conferences and text books and the reader is referred to the broader literature for more detail.
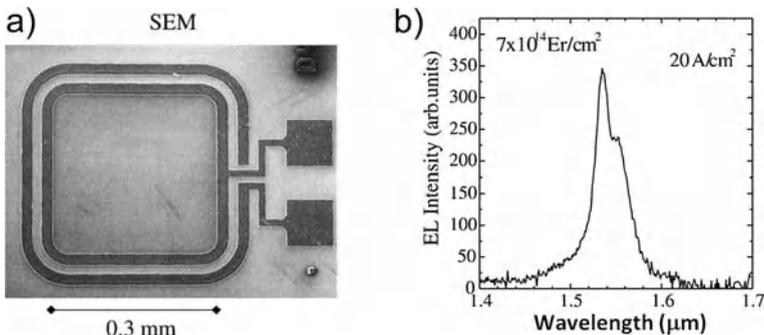
Fig. 16 (**a**) Scanning electron micrograph of a light emitting diode structure based on silicon-nanocrystal excitation of erbium. (**b**) Electrically stimulated Er emission from the device. The nanocrystals were formed from silicon-rich oxide and Er was introduced into the device by ion implantation (From [41])

Suffice it to say that the strong room temperature luminescence from silicon nanocrystals and the coupling to Er have been exploited to make prototypic devices such as the light-emitting diode shown in Fig. 16 [41], and to investigate all-optical amplifiers based on Er and Er/nanocrystal structures [35].

2.4 Optical Structures

The fabrication of 2D and 3D material structures is an important aspect of modern photonics and ion beams have found significant applications in this area. Ion beam patterning of materials is generally based on: a) direct removal of material by sputtering (e.g., focused ion beams [FIB]) or chemically assisted sputtering (e.g., reactive ion etching [RIE]), and b) ion beam modification of the etching rate of the materials. Here we present three novel approaches to nanostructure fabrication that serve to illustrate these approaches.

The first example, shown in Fig. 17a), illustrates the use of different reactive ion chemistries for the fabrication of a Au nanodisk array [42]. In this case, a gold film of 33 nm thickness was deposited onto a glass substrate, using a 3 nm Cr layer to increase adhesion. The surface was then coated with a self-assembled monolayer of polystyrene spheres before being subjected to reactive ion etching in O_2 and Ar plasmas to reduce the sphere dimensions and etch the underlying film. This provides accurate control over the final nanodisk size-distribution and was used to study the effect of particle size on the surface plasmon resonance, and to assess the dependence of the resonance on local refractive index changes induced by fluid flow within integrated microfluidic channels.

The second example, shown in Fig. 17b), shows a suspended beam structure produced in single crystal diamond by a combination of high-energy ion irradiation and FIB machining [33]. Here, MeV He ion implantation was used to create a narrow
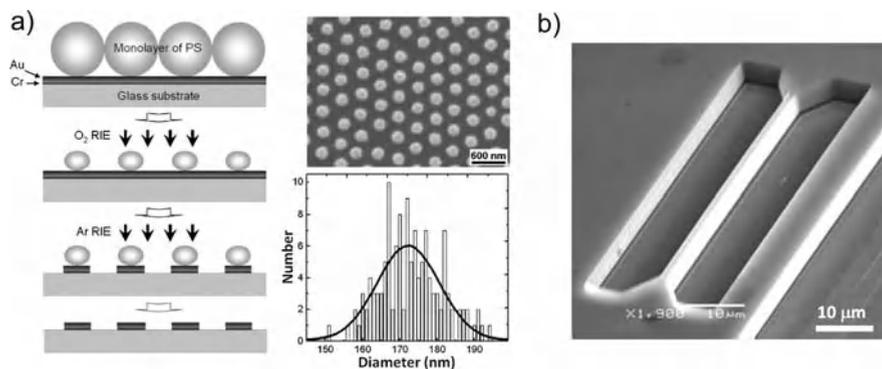
Fig. 17 (**a**) Schematic of the fabrication process, SEM image of the final Au nanodisks, and Measured nanodisk size distribution (From [42]). (**b**) SEM image of a bridge structure machined from single crystal diamond (From [43])

defect band deep in the diamond (3.5 μm for 2 MeV He) that acted as a sacrificial layer during subsequent etching. A FIB was used to open etching apertures between the surface and the buried damage layer and to undertake lateral patterning of the structure. A gap of 400 nm exists under the bridge because of the etching of the damage layer. Such nanoscale machining is being combined with ion implantation–induced color centers in diamond for photonic and quantum computing applications [33].

The final example of ion beam patterning is based on changes in the electrochemical etching rate of irradiated silicon [44]. In this case, etching is used to produce porous silicon, which has a porosity (and refractive index) controlled by the local hole concentration. Ion irradiation of p-doped Si wafers reduces the local hole concentration by creating traps and recombination centers in the material and thereby affects the subsequent growth of porous silicon. By undertaking controlled irradiation with a focused proton beam, it is therefore possible to define regions of different porosity. Figure 18 shows a dramatic example of such treatment, in which the lateral control of the irradiation fluence has been used to render images of different porosity. These images are formed by reflection of white light but porous silicon also exhibits strong room temperature luminescence. As the emission wavelength varies with porosity the images are also expected to be visible under luminescence imaging conditions. Clearly such control can be used to "write" complex photonic structures in silicon.
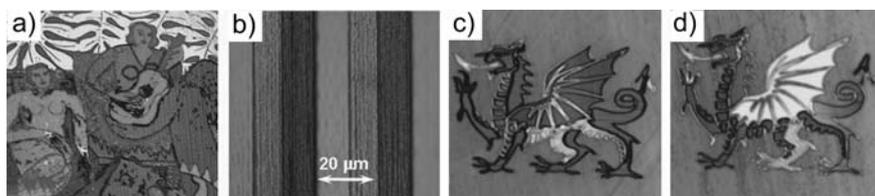
Fig. 18 Optical reflection images recorded under white light illumination: (**a**) the painting La Musique by Henri Matisse defined in a $500 \times 500 \, \mu m^2$ area; (**b**) *vertical lines*, each 10 μm wide; (**c**) and (**d**) dragon images corresponding to different irradiated doses, each in an area of $500 \times 500 \, \mu m^2$ (From [44])

3 Magnetic Thin Film Structures

Ion-irradiation offers many approaches to modifying the magnetic properties of thin films [45–47], including chemical doping, the synthesis of magnetic phases, layer intermixing, and flux pinning. The first of these is commonly employed for the synthesis of dilute magnetic semiconductors in spin-based electronics (spintronics) and includes extensive work on Mn-doped III-V semiconductors (e.g., GaP, InAs, GaAs) [12]. Other important applications of ion beams not covered here include ion beam assisted deposition of magnetic thin films and the use of ion cluster beams for atomic-scale smoothing of thin films and interfaces [48, 49].

3.1 Recording Media

An ideal material for high-density magnetic storage would consist of a thin film containing discrete, nanometer-scale, magnetically isolated, single-domain ferromagnets whose coercivity, size, orientation, and spatial distribution could be controlled by simple processing. The material should also have a smooth surface and a uniform refractive index. The former is critical in applications where the read/write head travels a few nanometers above the film surface; and the latter is important in magneto-optic applications to avoid beam diffraction and interference effects during the optical read/write processes. Materials synthesized or modified by ion beams meet many of these demands.

3.1.1 Heterogeneous Phases

Ion implantation can be used to make metallic nanocrystals within a broad range of host materials and has been shown to be an effective means of making films suitable for high density magnetic storage [50–53]. Figure 19 compares the microstructure and magnetic response of Co and CoPt nanocrystals formed by ion beam synthesis in sapphire [51].

   Implantation with Pt increased the coercivity and remnant magnetization of the Co nanocrystals, just as alloying with Pt does for Co films. However, ion implantation with Xe ions also produced an increase in coercivity and remnant magnetization, suggesting that changes in nanocrystal morphology (also expected from the Xe irradiation) and radiation damage also contribute to the observed magnetic response.
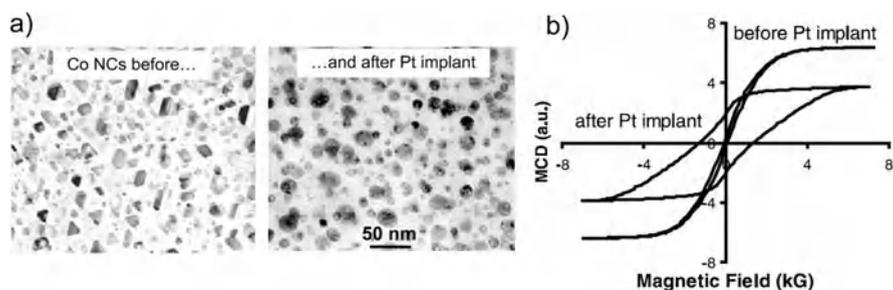
Fig. 19 (**a**) Transmission electron micrographs of Co and CoPt nanocrystals produced in sapphire by ion beam synthesis. (**b**) Magnetic hysteresis curves for the two samples shown (From [51])

3.1.2 Patterned Structures

Ion irradiation can affect the local magnetic properties of a material by changing its chemical ordering or crystallographic phase. For example, $Fe_{60}Al_{40}$ is paramagnetic in the chemically ordered B_2 phase but can be transformed by ion irradiation to the chemically ordered bcc phase that is ferromagnetic. Figure 20 shows a pattern written in such materials using a FIB [46]. By using suitable implantation masks,
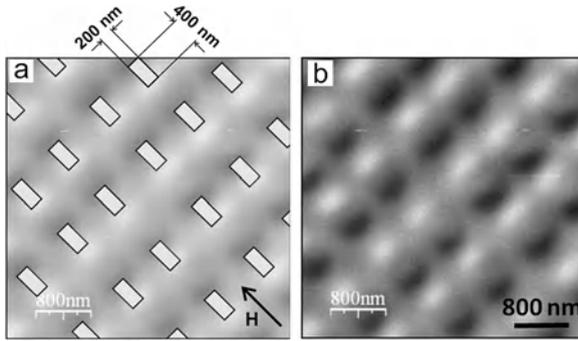
Fig. 20 (**a**) Schematic of the irradiation aperture showing rectangular areas of $400 \times 200\,nm^2$ and the applied field direction. (**b**) Magnetic force image of an irradiated film showing the stray field distribution (From [46])

sub-50 nm structures have also been defined in magnetic thin films by light ion irradiation [54].

3.2 Thin Film Devices

The magnetic properties of ultrathin films and multilayers depend on their surface and interface structure as well their chemical composition and atomic ordering (e.g., crystal structure). It is not surprising therefore that ion irradiation can influence such properties.

3.2.1 Magnetic Anisotropy

The ability to control the magnetic alignment of thin film structures is critical in applications as diverse as magnetic recording, solid-state compasses, automotive sensors, and nonvolatile magnetic memory. Perpendicular magnetic anisotropy, in which the magnetization is perpendicular to the plane of the thin film, is related to asymmetry in the atomic environment of a film and can be caused by atomic ordering, lattice strain, or the presence of interfaces. In many systems the interface anisotropy between two films is strong enough to overcome other contributions and can lead to perpendicular magnetization of the system. The Pt-Co thin film system is particularly well studied in this regard.

Light ion irradiation of thin film structures can produce atomic intermixing at the interface, and thereby increase the magnetic anisotropy of the structure, whilst maintaining the crystal structure of the films [55]. For Pt/Co/Pt, irradiation results in a continuous reduction in the room-temperature coercivity, out-of-plane anisotropy, and Curie temperature, with the system eventually undergoing a ferromagnetic-to-paramagnetic transition at high irradiation fluences, as illustrated in Fig. 21.
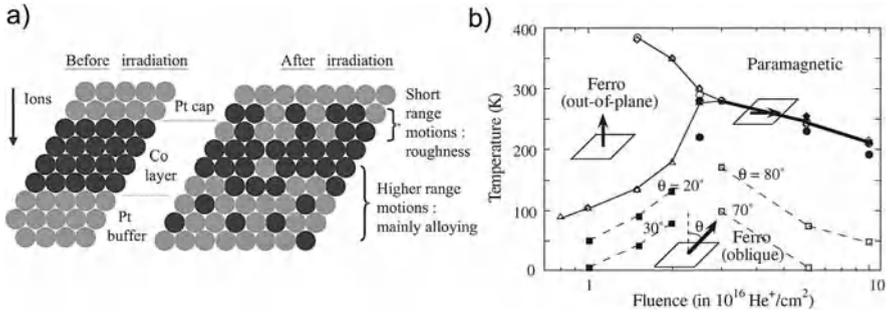
Fig. 21 (**a**) Schematic of a Pt/Co/Pt sandwich before and after irradiation (From [46]). (**b**) Temperature-fluence phase diagram for a Pt/Co/Pt sandwich irradiated with 30 keV He$^+$ ions (From [55])

These changes in magnetic properties are a direct consequence of ion beam induced interface roughening, lattice strain, and compositional changes, and serve to illustrate the potential of ion irradiation for tailoring such properties.

3.2.2 Exchange Coupling

Exchange coupling between magnetic thin films, in which the magnetization of one film is influenced by the proximity of another, can be exploited to make spin-dependent electronic devices, such as giant magnetoresistance (GMR) recording heads and spin valves. Exchange coupling can occur between two films in direct contact or between films separated by an intermediary layer. The interaction depends sensitively on the interface structure, and in the latter case, on the thickness of the spacer layer. Systems with exchange bias can also be fabricated from ferromagnetic and antiferromagnetic layers by freezing the spin alignment in the antiferromagnetic layer such that it produces an internal field that modifies the hysteresis response of the coupled ferromagnetic layer.

Ion irradiation can be used to modify the exchange coupling between layers by intermixing the interfaces, and can change the magnitude and orientation of the exchange bias field, thereby offering an important tool for device optimization [47].

3.3 Spintronics

Spintronic devices exploit the spin of charge carriers to control the injection, transfer, and detection of carriers and thereby extend the functionality of existing electronic devices. Since the magnetic properties of ferromagnetic semiconductors vary with carrier concentration, it is also possible to control the magnetic properties of these materials electrically or optically [12]. However, practical applications require materials with ferromagnetic ordering at room temperature. Figure 22 shows predicted and experimental Curie temperatures for a range of common semiconductor systems. Ion implantation has proven to be a very effective means of introducing
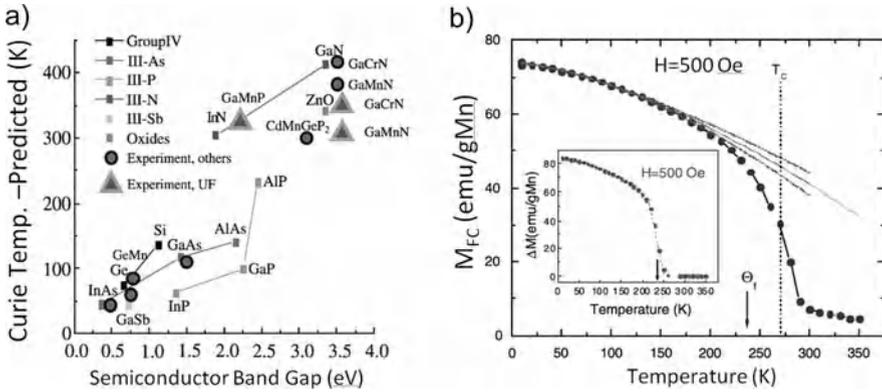
Fig. 22 (**a**) Predicted Curie temperature as a function of bandgap together with experimentally determined values. (**b**) Temperature dependent magnetization of a C-doped GaP sample implanted with ∼ 6% Mn, showing a Curie temperature of ∼ 270 K (From [12])

magnetic impurities, such as Mn, Fe, and Ni, into a variety of substrate materials, including GaN, SiC, and GaP [12]. This has proven to be an effective means of surveying material properties and of fabricating device structures. For example, Fig. 22(b) shows the temperature dependent magnetization of a Mn implanted GaP sample with a Curie temperature of ∼ 270 K.

4 Mechanical Structures

Ion beams are used extensively for modifying the mechanical properties of materials, with particular emphasis on the modification of surface properties for improved hardness and wear [56] and thin film and biological cell adhesion [57]. Most micro-electro-mechanical systems (MEMS) also make extensive use of ion beams, including techniques such as ion beam assisted deposition of thin films, reactive ion etching, focused ion beam etching, and ion beam lithography. Here we consider three novel applications, namely the fabrication of nanoporous membranes, the modification of mechanical properties, and the control of stress.

4.1 Nanopores Produced by Swift Heavy Ion Tracks

Swift heavy ions (SHI) with energies in the range 1–10 MeV per nucleon have very high electronic energy loss (or linear energy transfer [LET]) and can create latent tracks in many materials [58]. These tracks typically have diameters of a few nanometers and lengths > 10 μm, and exhibit preferential etching when exposed to reactive ambients. This can be used to produce nanoporous membranes or films that have direct device applications, or to produce structures that can be further modified (e.g., filling the pores with metal or semiconducting materials) to produce functional devices.
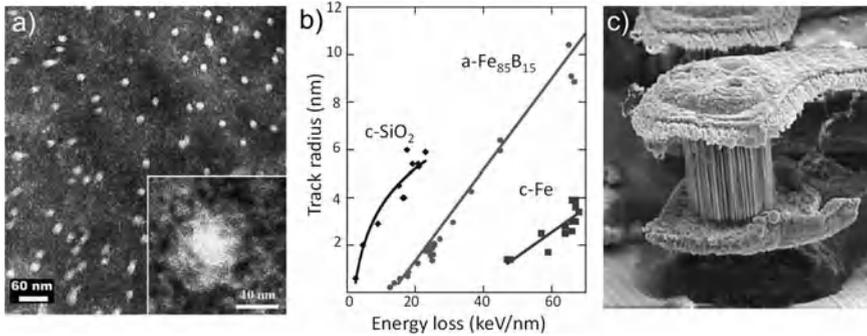
Fig. 23 (**a**) Transmission electron micrograph of individual ion tracks produced in polyimide by 11.1 MeV per nucleon Pb ions; (**b**) Track radius as a function of LET for different materials. The *solid lines* are calculations from a thermal spike model [58]. (**c**) Scanning electron micrograph of a wire array between two electrodes, produced by etching tracks and filling them with metal before etching the film (From [58])

Figure 23 shows discrete pores produced in polyimide by 11.1 MeV per nucleon Pb ions, together with a plot of pore size versus LET for different materials [58]. The final panel in this figure shows metal wires produced by etching the pores and filling them with metal [58]. The wires connect two electrodes and have been revealed by etching away the original thin film.

Similar structures exploiting conducting tracks or etched tracks incorporating metallic, semiconducting, or ionic matter have shown interesting electronic properties, including rectifying behavior, negative differential resistance, current oscillation, and self pulsing [59]. This is discussed at more length in the chapter Micro- and Nanoengineering with ion Tracks.

4.2 Mechanical Response

Nanoscale silicon structures find application in a broad range of micro-electromechanical devices, including high-frequency mechanical resonators, optical mirrors [60], and probes for resonance force microscopy. In all such applications the mechanical properties of silicon at the nanometer scale are of prime importance. The mechanical response of silicon components can be modified by alloying or coating the structures, but ion implantation also offers a very effective approach to such control.

As an example, Fig. 24 shows the effect of irradiation on the mechanical properties of Si cantilevers with dimensions $\sim$ 3 μm long, 250 nm wide, and 193 nm thick [61]. These cantilevers were ion implanted with either 100 or 35 keV Si- ions to render all or part of the structure amorphous. Their mechanical properties were then measured using an atomic force microscope (AFM), as shown in Fig. 24. Fully amorphized structures showed a 20% reduction, and partially amorphized structures an 11% reduction, in Young's modulus when compared to unimplanted cantilevers.
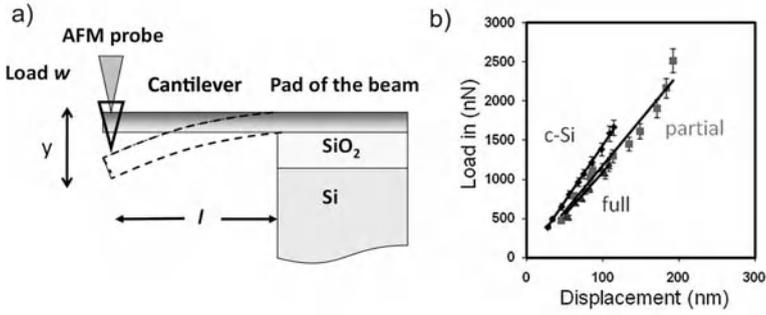
Fig. 24 (**a**) Schematic of the AFM technique employed to measure the mechanical response of cantilevers. (**b**) Load-displacement curves measured for unirradiated and irradiated structures (After [61])

The ability to induce such changes allows the response of MEMS components to be tuned for particular applications.

4.3 Stress Control

Unwanted compressive or tensile stresses are often produced in self-supporting thin film structures during the fabrication of MEMS devices. Ion implantation has been shown to be an effective means of controlling such stress and of reducing the temperature dependence of stress-induced curvature.

As an example of such applications, Fig. 25 shows an electrostatically actuated micromirror that forms one element of a steerable mirror array [60]. The polycrystalline silicon mirror is coated with 5 nm of titanium and 35 nm of gold to enhance its reflectivity. Implantation of these structures with 40 keV Si ions changed the
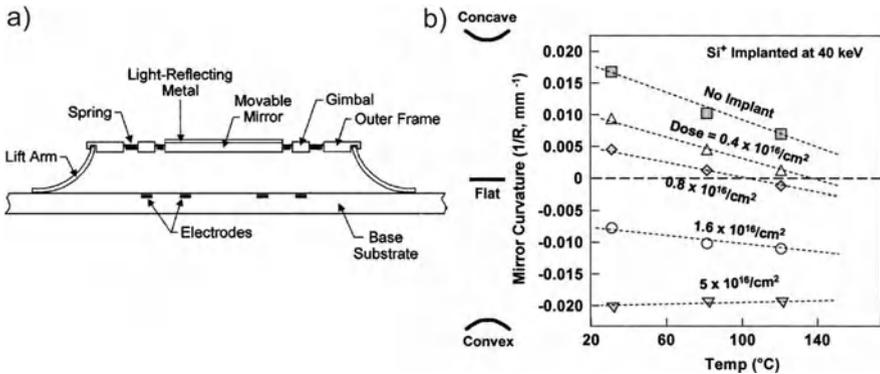
Fig. 25 (**a**) Schematic of a MEMS mirror device. (**b**) The measured mirror curvature as a function of irradiation fluence and temperature (After [60])

stress from tensile to compressive and reduced the temperature dependence of the curvature to zero at an optimum implant fluence.

Importantly, this technique has general applicability and can be applied to a broad range of films employed for MEMS applications, including Si_3N_4, SiO_2, SiC, diamond, and other metals and alloys.

4.4 Nanowire Catalyst

Nanometer-sized metal particles are used to catalyze the growth of inorganic nanowires. These particles absorb reactive species and act as a transport medium for the diffusion of reactants to the growing nanowire. The catalyst particles can be formed by chemical synthesis or the decomposition of a deposited thin film. However, recent work has shown that ion implantation is a very effective alternative to these processes in particular material systems. For example, Sood et al. [62] implanted silicon wafers with Pd to catalyze the selected area growth of silica nanowires. In this case, the implanted species precipitates on the sample surface, often in the form of a compound or eutectic phase, and acts to catalyze nanowire growth, as illustrated schematically in Fig. 26. An example of the resulting silica nanowires is also shown in Fig. 26. Such nanostructures are of interest for a broad range of applications, including as a high surface area substrate in biosensors [63].
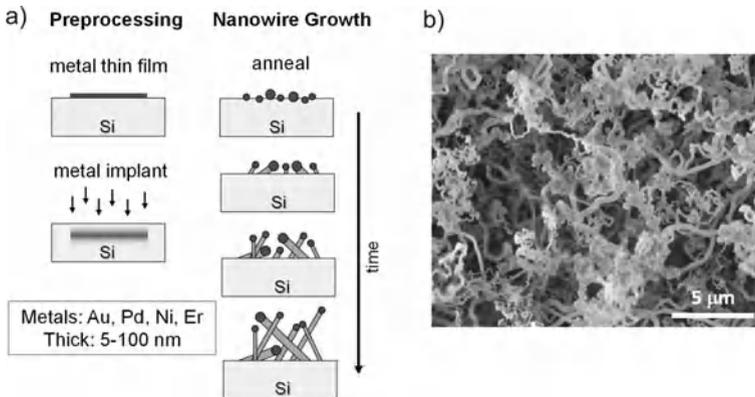
Fig. 26 (**a**) Schematic of the catalyst forming process. (**b**) Silica nanowires produced by annealing Au implanted silicon wafers

References

1. Chason E, et al. (1997) Ion beams in silicon processing and characterization. J. Appl. Phys. 81: 6513–6561
2. Keyes RW (1975) Effect of randomness in distribution of impurity ions on FET thresholds in integrated electronics. IEEE J. Solid-State Circuit 10: 245–247

3. Shinada T, et al. (2005) Enhancing semiconductor device performance using ordered dopant arrays. Nature 437: 1128–1131
4. Pok W, et al. (2007) Electrical characterization of ordered SiT dopant arrays. IEEE Trans. Nanotechnol. 6: 213–217
5. Kane BE (1998) A silicon-based nuclear spin quantum computer. Nature 393: 133–137
6. Schenkel T, et al. (2006) Strategies for integration of donor electron spin qubits in silicon. Microelectron. Eng. 83: 1814–1817
7. Jamieson DN, et al. (2006) Quantum effects in ion implanted devices. Nucl. Instrum. Methods B 249: 221–225
8. Hudson FE, et al. (2006) Coulomb blockade in a nanoscale phosphorus-in-silicon island. Microelectron. Eng. 83: 1809–1813
9. Marsh JH (1993) Quantun well intermixing. Semicond. Sci. Technol. 8: 1136–1155
10. Charbonneau S, et al. (1995) Quantum-well intermixing for optoelectronic integration using high-energy ion-implantation. J. Appl. Phys. 78: 3697–3705
11. Jiang H and Elliman RG (1996) Electrical properties of GeSi surface- and buried-channel p-MOSFET's fabricated by Ge implantation. IEEE Trans. Electron Dev. 43: 97–103
12. Pearton SJ, et al. (2003) Advances in wide bandgap materials for semiconductor spintronics. Mater. Sci. Eng. R-Rep. 40: 137–168
13. Mantl S (1994) Materials aspects of ion-beam synthesis of epitaxial silicides. Nucl. Instrum. Methods B 84: 127–134
14. White AE, et al. (1989) Mesotaxy-synthesis of buried single-crystal silicide layers by ion implantation. Nucl. Instrum. Methods B 39: 253–258
15. Kahng D and Sze SM (1967) A floating gate and its application to memory devices. Bell Syst. Tech. J. 46: 1288
16. Pavan P, et al. (1997) Flash memory cells – An overview. Proc. IEEE 85: 1248–1271
17. Hori T, et al. (1992) A MOSFET with Si-implanted gate-SiO_2 insulator for nonvolatile memory applications. IEEE IEDM Tech. Digest 469–472
18. Tiwari S, et al. (1996) A silicon nanocrystals based memory. Appl. Phys. Lett. 68: 1377–1379
19. Myers SM, et al. (1995) Chemical and electrical properties of cavities in silicon and germanium. Nucl. Instrum. Methods B 106: 379–385
20. Wongleung J, et al. (1995) Gettering of copper to hydrogen-induced cavities in silicon. Appl. Phys. Lett. 66: 1231–1233
21. Chen F, Wang XL, and Wang KM (2007) Development of ion-implanted optical waveguides in optical materials: A review. Opt. Mater. 29: 1523–1542
22. Beghoul MR, et al. (2007) Photonic band gap grating in He+-implanted lithium niobate waveguides. Opt. Quantum Electron. 39: 333–340
23. Bohren CF and Huffman DR (1983) Absorption and Scattering of Light by Small Particles. New York: John Wiley and Sons
24. Shalaev VM (1996) Electromagnetic properties of small-particle composites. Phys. Rep. Rev. Sec. Phys. Lett. 272: 61–137
25. Flytzanis C, et al. (1991) Nonlinear optics in composite materials 1: Semiconductor and metal crystallites in dielectrics. Prog. Opt. 29: 321–411
26. Buchal C, et al. (1994) Ion-implantation of optical materials. Annu. Rev. Mater. Sci. 24: 125–157
27. Fu L, et al. (2001) Tuning of detection wavelength of quantum-well infrared photodetectors by quantum-well intermixing. Infrared Phys. Technol. 42: 171–175
28. Keller U and Tropper AC (2006) Passively modelocked surface-emitting semiconductor lasers. Phys. Rep. Rev. Sec. Phys. Lett. 429: 67–120
29. Keller U, et al. (1996) Semiconductor saturable absorber mirrors (SESAM's) for femtosecond to nanosecond pulse generation in solid-state lasers. IEEE J. Sel. Top. Quantum Electron. 2: 435–453
30. Lederer MJ, et al. (2001) Ion-implanted InGaAs single quantum well semiconductor saturable absorber mirrors for passive mode-locking. J. Phys. D-Appl. Phys. 34: 2455–2464

31. Zarrabi JH, Portnoi EL, and Chelnokov AV (1991) Passive-mode locking of a multistripe single quantum-well Gaas-Laser diode with an intracavity saturable absorber. Appl. Phys. Lett. 59: 1526–1528

32. Ng WL, et al. (2001) An efficient room-temperature silicon-based light-emitting diode. Nature 410: 192–194

33. Warldermann FC, et al. (2007) Creating diamond color centers for quantum optical applications. Diam. Relat. Mat. 16: 1887–1895

34. Kippenberg TJ, et al. (2006) Demonstration of an erbium-doped microdisk laser on a silicon chip. Phys. Rev. A 74: 051802: 1–4

35. Lourenco MA, Gwilliam RM, and Homewood KP (2007) Extraordinary optical gain from silicon implanted with erbium. Appl. Phys. Lett. 91: 141122: 1–3

36. Leong D, et al. (1997) A silicon/iron-disilicide light-emitting diode operating at a wavelength of $1.5\,\mu m$. Nature 387: 686–688

37. Iyer SS and Xie YH (1993) Light-emission from silicon. Science 260: 40–46

38. Canham LT (1990) Silicon quantum wire array fabrication by electrochemical and chemical dissolution of wafers. Appl. Phys. Lett. 57: 1046–1048

39. Takeda H, et al. (1990) Quantum size effects on photoluminescence in ultrafine Si particles. Appl. Phys. Lett. 56: 2379–2380

40. Pavesi L, et al. (2000) Optical gain in silicon nanocrystals. Nature 408: 440–444

41. Pacifici D, et al. (2003) Er doped Si nanostructures. Mater. Sci. Eng. B-Solid State Mater. Adv. Technol. 105: 197–204

42. Zheng YB, et al. (2008) Systematic investigation of localized surface plasmon resonance of long-range ordered Au nanodisk arrays. J. Appl. Phys. 103: 014308: 1–9

43. Olivero P, et al. (2005) Ion-beam-assisted lift-off technique for three-dimensional micromachining of freestanding single-crystal diamond. Adv. Mater. 17: 2427–2430

44. Watt F, et al. (2007) Proton beam writing. Mater. Today 10: 20–29

45. Chappert C, et al. (1998) Planar patterned magnetic media obtained by ion irradiation. Science 280: 1919–1922

46. Fassbender J and McCord J (2008) Magnetic patterning by means of ion irradiation and implantation. J. Magn. Magn. Mater. 320: 579–596

47. Fassbender J, Ravelosona D, and Samson Y (2004) Tailoring magnetism by light-ion irradiation. J. Phys. D-Appl. Phys. 37: R179–R196

48. Yamada I and Toyoda N (2005) Summary of recent research on gas cluster ion beam process technology. Nucl. Instrum. Methods B 232: 195–199

49. Fenner DB, et al. (2001) Surface processing with gas-cluster ions to improve giant magnetoresistance films. J. Vac. Sci. Technol. A-Vac. Surf. Films 19: 1207–1212

50. Budai JD, et al. (1997) Controlling the size, structure and orientation of semiconductor nanocrystals using metastable phase recrystallization. Nature 390: 384–386

51. Meldrum A, Boatner LA, and White CW (2001) Nanocomposites formed by ion implantation: Recent developments and future opportunities. Nucl. Instrum. Methods B 178: 7–16

52. Withrow SP, et al. (2003) Ion beam synthesis of magnetic Co-Pt alloys in Al_2O_3. J. Magn. Magn. Mater. 260: 319–329

53. Zhou SQ, et al. (2008) Crystallographically oriented Co and Ni nanocrystals inside ZnO formed by ion implantation and postannealing. Phys. Rev. B 77: 035209: 1–12

54. Devolder T, et al. (1999) Sub-50 nm planar magnetic nanostructures fabricated by ion irradiation. Appl. Phys. Lett. 74: 3383–3385

55. Ferre J, et al. (2003) Magnetic phase diagrams of He ion-irradiated Pt/Co/Pt ultrathin films. J. Phys. D-Appl. Phys. 36: 3103–3108

56. Pelletier J and Anders A (2005) Plasma-blased ion implantation and deposition: A review of physics, technology, and applications. IEEE Trans. Plasma Sci. 33: 1944–1959

57. Liu XY, Chu PK, and Ding CX (2004) Surface modification of titanium, titanium alloys, and related materials for biomedical applications. Mater. Sci. Eng. R-Rep. 47: 49–121

58. Toulemonde M, et al. (2004) Track formation and fabrication of nanostructures with MeV-ion beams. Nucl. Instrum. Methods B 216: 1–8

59. Fink D, et al. (2007) Swift heavy ion irradiation as a tool for creating novel nanoelectronic structures. Radiat. Eff. Defects Solids 162: 543–555
60. Jin S, et al. (2003) Control of microelectromechanical systems membrane curvature by silicon ion implantation. Appl. Phys. Lett. 83: 2321–2323
61. Virwani KR, et al. (2004) Modification of mechanical properties of silicon nanocantilevers by self-ion implantation. Appl. Phys. Lett. 84: 3148–3150
62. Sood DK, Sekhar PK, and Bhansali S (2006) Ion implantation based selective synthesis of silica nanowires on silicon wafers. Appl. Phys. Lett. 88: 143110
63. Sekhar PK, Ramgir NS, and Bhansali S (2008) Metal-decorated silica nanowires: An active surface-enhanced Raman substrate for cancer biomarker detection. J. Phys. Chem. C 112: 1729–1734

Luminescence, Ion Implantation, and Nanoparticles

Peter Townsend

1 Introduction

Ion implantation offers control of dopants, defect structures, and crystallinity of surface layers. From the viewpoint of luminescence and laser properties, the ion beams both change and excite the near surface material. The changes can be considered for at least the following range of applications.

1. The first is to use the luminescence as a diagnostic tool of the host material, both in its initial and in later modified forms (as discussed in Box 7).
2. The second approach is to introduce dopants, such as rare earth ions, which offer efficient luminescence in later applications. This step may also result in laser action.
3. A third option is to use the implants and/or the lattice damage associated with implantation to define waveguide structures, which are required for the optics of waveguide lasers, second harmonic generation, upconversion, etc.
4. A fourth opportunity exists when the implanted dopants, or defect structures, result in the creation of nanoparticles within the surface layer.

Nanoparticles, phase precipitates, and extended defect structures all offer benefits in providing new optical and luminescence properties. These various luminescence related facets of ion beam modification require different types of optimisation. For example, the precipitation of rare earth dopants into nanoparticles can minimise the luminescence intensity in a waveguide laser, whereas agglomerating metal precipitates into nanoparticles can result in highly nonlinear interactions that are required for photonic switching. Thermal treatments can potentially define the requisite states.

This chapter discusses some of these various opportunities and includes examples of unexpected results, indicating that the surface implants can modify the substrate crystallinity and bulk optical properties. The relevance for laser devices and

P. Townsend (✉)
Science and Technology, University of Sussex, Brighton BN1 9QH, UK
e-mail: p.d.townsend@sussex.ac.uk

modifications of thin films, such as optical components on the surface of semiconductor hosts, are mentioned. Some broad considerations are offered on the various processes which can occur. For combinations of implantation and luminescence studies, there exists a relatively limited literature. Since the topic is in some aspects at an early stage of development, it has resulted in conflicting views and/or insufficient data to fully understand these complex processes. Despite this, even currently available examples show that one can offer new analytical techniques (see Box 7) which aid in the empirical development for forming valuable devices. Realistically, exploitation of the optical properties of nanoparticles formed by ion implantation is mostly in an empirical, rather than a predictive, phase.

2 Luminescence from Ion Implanted Material

Luminescence features of insulators and semiconductors result from ion beam implantation, not only in the immediate near surface regions where the ions penetrate, but also from a more extended volume. This is the result of charge movements, radiation enhanced ionic diffusion, and excitonic or other electronic defect formation processes. Exciton diffusion can extend over ranges as long as a μm, which may be far greater than the implant depth. Excitonic or ionisation types of damage routes are well documented for alkali and alkaline earth halides as well as organic materials, and there exist a few unequivocal examples of long-range effects (i.e., beyond the implant zone). The overall modifications in property can be sensed by luminescence. Unfortunately, many authors do not appreciate that changes can result from beyond the immediate implant region. In particular, it is less commonly recognised that optical effects can be modified by stresses generated in the implant zone. These stresses induce relaxations, ionic transport, and even phase transitions, which extend deep into the surface layers. Sensitive features such as the emission wavelength of rare earth ions exhibit wavelength and lifetime differences in zones of different stress. Long-range structural modifications have been studied in detail for $SrTiO_3$ and ZnO (as mentioned later). Both these materials can exist in a range of metastable phases which are sensitive to impurities and stoichiometry, and so they are responsive to stresses induced by the surface implants. Similar conditions exist in many insulators, and therefore one suspects that more careful analysis might reveal further long-range examples of implant induced change. Long-range events may have been overlooked because it has been fashionable to use high implant fluences to induce precipitate phases or inclusions of impurity compounds and/or nanoparticle production. A focus on these high fluence effects will normally fail to notice the more subtle, deeper layer changes. Again this is unfortunate, because in many cases, such as the use of Si implants in SiO_2 layers on silicon, the long-range features include charge transport beyond the implant zone, which in turn influences the luminescence and laser properties.

Luminescence information is available both during ion implantation and postimplantation. Subsequent excitation, as with cathodoluminescence (CL) or

photoluminescence (PL), is highly sensitive to the implant damage, stress, and the presence of "point" defects, as well as precipitates and nanoparticles [1]. One should note that in all cases, the concept of highly localised "point" defects was just a convenient early descriptive simplification. In reality, luminescence (or lasing) involves complex defect sites. These are frequently an aggregation of various intrinsic and extrinsic defects which form as the host material accommodates local charge equilibrium and minimises the overall lattice stress. The emission spectra and changes in excited state lifetime are also sensitive to quite long-range interactions, evidenced by over tens of nearest neighbour distances [2]. Hence ion implanted laser waveguide materials will show changes in emission wavelength and lifetime which can include a range of values across the wavelengths used for the excitation. Such analysis is often referred to as site-selective spectroscopy.

References to articles which have discussed these concepts span more than 40 years (as exemplified by other chapters in this current book), but a search of the more recent literature that specifically involves nanoparticles (e.g., by key words via, say, Google Scholar) generates just a few hundred references with a very limited range of materials. There is particular emphasis on examples of data for Si implants in silica and silicon oxide films (e.g., [3–10]), some examples of Er implants for infra-red fibre optic communication lasing, examples of semiconductor impurity nanoparticle formation [11–14], examples with ZnO formation [15, 16], and a quite limited reference to other studies such as Er in AlN [17], etc. The only general patterns which emerge are that annealing treatments can alter the luminescence spectra and intensity. For the Si nanoparticles, thermal processing shifts the emission band. In some cases, authors have suggested that wavelength shifts are linked to quantum size effects for the nanoparticles. However, the nanoparticles generally exist over a range of sizes, so it is unclear how to separate emission band broadening from quantum effects rather than just a different range of defect sites and long-range interactions. As already mentioned, the excitation of photoluminescence often involves a deeper thickness of material than the implantation range. Modelling of luminescence sites, including modified intrinsic defects, is difficult, because it is unclear how to separate intrinsic signals from the true nanoparticle signals. Further, some authors maintain that the nanoparticles often include oxide or other barrier regions, or even that doping of the Si nanoparticles (e.g., with Er) may be primarily on the surface of the silicon. Such conflicts of interpretation are inevitable with the information that is available, but for advances in technology it is not a serious problem. Indeed, in reality much of condensed matter technology has progressed effectively via intuitive empiricism. The demonstrated advantages of high fluence implants means this area of experimentation will be an ongoing and expanding field.

3   Examples of Luminescence from Ion Implanted Systems

Before addressing nanoparticle data, it is instructive to commence with a summary of luminescence work for low implant fluences where phase precipitation or nanoparticle production is unimportant. The first examples of implanting insulators

with rare earth ions for luminescence studies were made more than 40 years ago, as detailed in a book describing the earlier work, Box 7, and [18], but the number of more recent publications is still quite limited. Rare earth ions are ideal for implantation because they often have high luminescence efficiency and are readily detected via their line emission. In general, the lattice damage associated with the implantation seriously detracts from the overall efficiency and some postimplant annealing is required. For many oxides, such as alumina, silica, or zinc oxide, the luminescence signals increase with annealing above about $800\,^{\circ}$C. This is the temperature range in which intrinsic defects formed by the implantation are removed. Higher temperature anneals can be slightly more effective (e.g., $\sim 1000\,^{\circ}$C) in some oxide materials, but even higher temperature processing at, say, $1200\,^{\circ}$C is usually counterproductive because it often reduces the emission intensity. The fundamental problem of the higher temperatures is that the impurity ions become mobile, and therefore they aggregate into small clusters (or nanoparticles). The light intensity is typically increased by the presence of individual well separated ions, so pairs, clusters, or particles reduce the luminescence efficiency. Similarly, the furnace treatments are most effective for low fluence implants ($\sim 10^{14}$ ions/cm^2). However, for applications such as the production of rare earth waveguide lasers, it is essential to inject high fluences of the active laser ion (10^{16}–10^{17} ions cm^{-2}). At these fluences simple furnace annealing methods result neither in the required improvements in the signal intensity, nor in the excited state lifetimes, as demonstrated by Polman [3]. He showed that at high Er impurity concentration in silica or alumina, the dopant ions are not accommodated as dispersed individual ions within the host structure, but instead precipitate out either as clusters of rare earth ions or in a rare earth oxide phase. An alternative annealing strategy used by Townsend and coworkers was to use pulsed laser annealing [4, 19]. In their approach, the laser energy deposition rate was sufficiently high that the clusters in the implant zone were successfully dissolved. One must note that pulsed laser heating not only gives a rapid temperature rise, but also a rapid cooling. Thus with short pulse anneals, the rapid cooling causes freezing into a dispersed state of the impurities. For the rare earth ions that have moved sufficiently far apart, they return to a longer excited state lifetime, and also are not subject to interactions which give rise to concentration quenching of the signals. The dispersed ions are thus locked in place and are not able to recluster. Since thermal anneals show effects above $\sim 800\,^{\circ}$C, one assumes that rapid cooling to below this temperature is the key factor for inhibiting diffusion. In the case of Eu implanted silica and sapphire, the pulsed laser anneals raised the luminescence efficiency by factors of almost 100 times even for high fluences, whereas optimal furnace treatments achieved only improvements of less than ~ 20 times [3, 4, 6].

The overall concept of using implantation to introduce luminescence sites in the target materials is tempting, since luminescence efficiency can be high and there is the ability to control the depth, ion fluence, and ion species. Therefore, in applications which require nonvisible bar code markers, or other security indicators, one can achieve the objectives with relatively small ion fluences. As an example, one can use a controlled mixture of different rare earth ions as a marker code for valuable items such as large gemstones (including diamond, sapphire, or ruby). The

ion fluences will be sufficiently small that visually the surfaces are unchanged but with PL or CL excitation one can read the identifier code even for very small ion beam fluences. The possibility of using more than one ion offers additional coding security.

One unusual example [20] employed Au implants in a nanopore cage structure, which for low fluences is assumed to accommodate the implant ions in empty cages and generates signals near 407 nm (3.05 eV). Increasing the fluences above 3×10^{16} ions cm^{-2} alters the damage and the Au sites, and results in Au nanoclusters which display different spectral responses. The low fluence data are particularly interesting because they imply that one could inject isolated phosphors at a high density but which would remain sufficiently screened from each other that they would fluoresce with a very high efficiency. Clearly this concept of a cage host should be pursued with other dopant ions.

The more general case of injecting high fluences of active phosphor ions uniformly throughout a deep layer of μm dimensions is much more problematic. For example, using Er ions for a near infrared waveguide laser requires a high density of sites. For the relatively high mass Er the implanter must have MeV capability. Such prolonged high fluence implants both raise the production costs as well as requiring techniques to remove the radiation damage and optimise lasing performance. On the positive side, one can combine ions, targets, and/or dopant levels, which are not normally compatible because the implantation route does not have the classical constraints of thermodynamic solubility of ions into the host target. One could therefore exploit this approach for, say, rare earth implants in quartz, which is a material which does not naturally allow a high impurity content of these ions. Indeed, there have been some studies of alkali or Ge ion implants in quartz to control the structure of the material [21, 22]. Problems include the inherent amorphisation of the lattice (to high density silica glass) and differences in recrystallisation associated with the implant ions. The work has been the subject of a recent review of luminescence from ion implanted quartz [23]. A quartz example [24] also emphasises an unexpected result, in that one can induce Si nanoparticles by subsequent Au implants.

Conventionally, the implantation is carried out, and subsequently furnace or laser annealing is used to regenerate the host lattice. Unfortunately, not all materials recover completely, or to the same phase, once they have been highly damaged or amorphised. A different strategy could be to use pulsed laser annealing *during* the implantation process so that at no time does the host reach an irreversible amorphised condition. Experimentally the technique is simple, and it has the advantage that one can control the base temperature of the substrate. The base temperature is actually quite critical; and, for example, in Sussex, for production of blue reflective car mirrors by control of the size of Ag nanoparticles, it proved essential to stabilise the substrate temperature within about 5 degrees, and precisely control the time dependence of the ion beam flux.

Laser anneal pulses within the implant cycle can also be introduced when the host material is at low temperature. This offers some unusual features because the thermal conductivity of the host can change significantly from the normal nominal room temperature implants, which in reality heat the surface significantly during

high fluence rate implants, such as those used to form nanoparticles. A particularly interesting example involved implantations in sapphire, since the thermal conductivity of the Al_2O_3 host rises during cooling and at 55 K exceeds that of copper. This change in thermal conductivity has a strong influence on defect retention, amorphisation, and, in the case of C implants, Al-C bond formation [25].

4 Luminescence Structures—Including Ion Implanted Waveguides

Ion implantation has proved to be a very versatile route for the formation of optical waveguides in an extremely wide range of materials [18, 26, 27]. The method is simple, in that with light ion implantation (e.g., He^+) the host material is primarily damaged at the end of the ion range. For crystalline targets, such as quartz, the lattice structure is damaged, and in extreme cases, is amorphised (in the quartz case to a high density silica). The net effect of the end-of-range damage is to generate a low refractive index layer below the normal (higher) index layer. Thus one has an optical waveguide. Damage within the crystalline waveguide can often be annealed out without losing the low index barrier. The technique has been demonstrated on more than 100 insulators. This combination of ion implantation and waveguide formation has numerous features relevant to luminescence and lasing, since one can commence with a bulk material which is already doped to provide lasing action. The implants merely confine light in the surface region, forming a waveguide laser. In doing so there is a decrease in the threshold power to drive the laser, which also results in very high slope efficiency. Various examples of such waveguide lasers have been demonstrated by several groups using examples such as Nd:YAG, other garnets, BGO, or Nd : $LiNbO_3$. Whilst in principle the waveguide is basically the original host material, the reality is that there are relaxations, perhaps caused by defects or stresses, which slightly modify the host lattice. For Nd:YAG this is apparent from the luminescence emission, which changes the relative intensity of the Nd emission lines and also can introduce some minor wavelength shifts. Changes in wavelength and transition line intensity merely indicate that the crystal field around the Nd has been altered. The material still forms a high grade waveguide laser, but has the unexpected advantage that lasing can be tuned to more transitions than were considered possible in Nd:YAG. When operated as a laser, the ion implanted guides often function at wavelengths displaced from those of the bulk, or, as for Nd:YAG, operate efficiently on alternative transitions. Changes in excited state lifetimes caused by stress or the presence of defects can be noted. A more serious potential problem would exist for extremely fast pulse operation, because lifetimes could differ in different regions of the guide, which would lead to pulse lengthening. This is a familiar problem with guides made by other production methods.

Similar waveguide confined structures can be used for second harmonic generation. Second harmonic generation in waveguides has the obvious advantage that a high power long wavelength (e.g., at 1064 nm) pump signal is confined within

the waveguide. Harmonic efficiency depends on the power density and length over which the interaction can be maintained. By suitable choice of guide design, the nonlinearity of the crystal waveguide produces the shorter wavelength harmonic (at 532 nm) travelling at the same velocity. Thus the continuous phase match of the two signals allows an efficient energy conversion to the harmonic. The phase match has also been achieved in ion implanted waveguides by coupling between a pair of adjacent waveguides. This design extends the short wavelength limit of harmonic generation. Similar guide structures allow up-conversion to wavelengths other than the harmonic. All these luminescence-type optical effects from ion implantation were demonstrated in the early 1990s [18], but further examples and developments with a commercial emphasis seem extremely slow to appear, despite the very high grade demonstration devices.

5 Benefits and Problems of Nanoparticle Formation

As already mentioned in the chapter "Nanomaterials Science with Radioactive Ion Beams" and Box 7, high fluence ion implantation can result in phase precipitation or nanoparticle formation [18, 28, 29], either of the host material or from the implants. For optical applications the presence of the nanoparticles can be either helpful or problematic. There have been numerous studies of metal implants in glass. In the case of copper ions in silica glass [30], subsequent luminescence studies indicate the existence of very many different narrow-line luminescence bands, which suggests that copper is bonded in the host in a range of different environments. The options span inclusion as isolated ions to large scale nanoparticles. Note that the high sensitivity of the luminescence measurements allows one to detect small quantities of isolated ions, or small scale defect clusters, even in the presence of large nanoparticles which could represent the bulk of the implanted copper. The large Cu nanoparticles certainly exist and are readily characterised by absorption and reflectivity data. There are therefore competing opportunities to form and exploit the small scale copper sites for luminescence or the large scale nanoparticles for nonlinear optics. Indeed, copper implants have resulted in demonstrations of extremely high coefficients of third-order nonlinearity [31].

The fact that there is a mixture of small and large scale inclusions resulting from a metallic implant is inevitable, since implantation is a statistical process with a range distribution. Large clusters form near the peak of the distribution, whereas small or isolated metal sites will occur in the tails of the depth distribution. In some cases, such as Ag or Cu in glass, the distribution of particle size resembles the range distribution. This is not always the case, and, for example, for Ga implants in glass at a certain temperature, there is a spectacular split between a few large Ga nanoparticles near the surface and a deeper layer with a host of very small particles [32]. The reasons for this sudden change in size distribution, and also the sensitivity to implant temperature, are unclear. However, two noteworthy factors are that Ga has a low melting point (29.8 °C) and the unusual feature of a 10%

increase in density on melting. More generally, for most metallic nanoparticles, the melting point is significantly suppressed below the bulk value for smaller particles (e.g., down as low as $\sim 400\,°C$ for Fe or Au). Nevertheless, even with only partial control of the nanoparticle sizes, optical applications have been demonstrated for Cu, Ag, and Ga implants. These range from the excellent third-order nonlinearity for Cu; a directional reflectivity feature (Cu); blue reflective car mirrors (Ag); to noncolour–distorting mirror sunglasses (Ga). Gallium ion implants are also being explored in an attempt to generate optical absorption in SiC for high density data storage applications. In principle, one might use the same implant approach, but sense the signals by luminescence.

Controlled implants to form nanoparticle compounds, such as ZnS, GaAs, or GaN, etc., in glass (e.g., [11–14]), pose the obvious problem of very large implant fluences, but nevertheless they have been attempted. Many examples have been referenced in a recent review [14] in which narrow lines were excited from excitonic relaxation in doped semiconductor nanoparticles. The signals originate at exciton states, and since exciton wave functions are quite extensive, the precise energy of the luminescence decay is strongly dependent on the size of the nanoparticle. Implants in silica to form ZnO [13] underline some of the difficulties in forming impurity compound nanoparticles by implantation. For the ZnO case, one must first implant the Zn before the O is added. Attempts to implant Zn and then anneal the material prior to the addition of excess oxygen merely result in Zn nanoparticles. For the ZnO the fluence, particle size, and annealing treatment result in movement of the absorption band edge (and hence changes in luminescence spectra). Such data almost inevitably disguise the fact that the nanoparticles are not of a single size, and so the absorption and emission signals are a convolution of a range of particle size effects. Similar responses were noted in later work [16], also using silica glass as the target. In this case, anneals to $600\,°C$ resulted in Zn nanoparticles throughout the implant layer, but on raising the temperature to $700\,°C$ in an oxygen atmosphere, there was metal movement towards the surface and these outer particles were converted to ZnO particles. These generated exciton luminescence at 375 nm, whilst the Zn defects emit near 500 nm. None of these effects are totally surprising, although they do indicate that there can be a variety of outcomes from nominally the same treatments, and in this case there were a mixture of ZnO particles at the surface and Zn particles deeper in the solid.

For ZnO nanoparticle formation in SiO_2, one might assume that oxygen could be released by atomic displacement that occurs in the host structure; but in more general terms, injection of element X prior to the implants of element Y may form nanoparticles of XY, whereas the reverse order could lead to a quite different chemical outcome. There is also a shortage of published information on the quality and purity of such nanoparticle inclusions. Indeed, even for metal implants in oxides, there are suggestions that the particles have outer layers which include oxygen. Compared with many of the luminescence applications, which only require dopants at the ppm level, the implanted nanoparticles inevitably require fluences several orders of magnitude larger. Consequently, there is much more extensive associated damage and destruction of the host and the nanocrystals. As mentioned above, the

normal practice is to postanneal the implanted material, but this can be too late and ineffective, so it may be desirable to implant in small fluence implant steps interspaced with anneal steps.

Once particle size, impurity content, and intrinsic defect features have been systematically controlled, then the nanoparticles can act as precise wavelength phosphors or laser materials. Although not immediately obvious, additional implants, or the inclusion of further impurities, or more than one type of nanoparticle, can modify the charge trapping within the layer as well as the precipitation process [8, 10, 12, 24]. A recent example [10] has reported that the luminescence efficiency of Er in a silica thin film is sensitive to Si nanoparticles, both under electroluminescence (EL) and photoluminescence. The EL signal intensity was decreased by first forming Si nanoparticles in the layer, whereas the PL signals were significantly increased. It is not yet clear if the annealing of the implants was fully optimised, not least because only furnace anneals were employed, whereas the earlier examples suggested there were greater benefits available from pulsed laser anneals.

One notes that there is great interest in attempting to combine luminescence and/or lasing with silicon technology, and this has spawned an extensive number of attempts to form Si and/or Er implants in oxide layers on a silicon substrate. Spectra differ with the type of excitation, e.g. from EL, CL, or PL. In most cases the authors assume that Si nanoparticle formation is a key factor in this process. There is general agreement that there is light emission near 700–800 nm, probably from the excess Si, or signals near $1.54\,\mu$m from Er. Many models involve defects, nanoparticle surfaces, charge trapping, alignment of particles, and the role of the initial oxide film; but there is no unanimous view as to the details, although some protocols of implantation and annealing generate more light intensity. As a brief example, Si nanoparticles formed in an SiO_2 thin film [7] were excited with argon laser light at 514 nm, and there were distinct differences in the emission spectra as a function of particle size, with particles of 2–5 nm diameter producing light near 900 nm, whereas small particles generated signals near 700 nm. Measurements of how the signal varied with depth showed that surfaces and defects were influencing the luminescence spectra and intensity. In contrast, a more recent article [9] reported luminescence with long wavelength signals, but related no wavelength shifts linked to the Si particle size. However, the authors recorded a broad emission feature centred near 820 nm, which shifted in peak value as a function of excitation power, and they noted some differences in luminescence lifetime linked to the particle sizes. These examples seem to be typical of the diversity of responses reported in the current literature, where a range of conditions exist for sample preparation and processing.

Totally different emission bands are noted in some cases, such as the blue luminescence signals associated with oxygen vacancies in most oxide materials. Such signals may also be linked to oxide layers on the nanoparticles. Separation of host and particle signals is not trivial.

Finally, there are just a few examples with evidence that nanoparticle formation in the surface layer triggers relaxations and phase events which extend deep into the bulk of the material. The examples observed so far have shown these

long-range effects via studies of luminescence. Data have been gathered mostly for $SrTiO_3$ [33–35]. More recently, luminescence data from ZnO has indicated similar long-range effects into the bulk region; and one assumes that many other examples will emerge for other materials, particularly if their structural stabilities are equally sensitive to pressure, impurities, or defects.

6 Conclusion

The preceding rather brief summary of examples of ion implantation related to luminescence effects is encouraging in that there have been varied and successful examples demonstrated. Unfortunately, there is not a strong overlap between the implantation, luminescence, and laser communities. Therefore the very obvious potential of implantation-related luminescence has had rather few practical demonstrations of applications. In order to optimise performance, many of the processes require a sound understanding of ion implantation, the structure and stability of defects and their annealing, as well as a clear appreciation of why ion beams are the best route to generate a particular device. Many of the current examples lack the in-depth development which is required to fully understand the processes. Therefore, the results merely indicate a first indication of performance. Such data can be quite misleading, as well as being inadequate scientifically. For example, if the initial results are poor, they may deter further study. The clear message is that the field has potential and can be greatly exploited if sufficient effort is expended; and it could result in both publications of the new understanding of the science as well as novel or improved commercial products.

References

1. Townsend PD, Karali T, Rowlands AP, Smith VA and Vazquez G, Mineralogical Magazine, 63, 211–226, 1999.
2. Townsend PD, Jazmati AK, Karali T, Maghrabi M, Raymond SG and Yang B, J. Phys. Condens. Matter **13**, 2211, 2001.
3. Polman A, Jacobson DC, Eaglesham DJ, Kistler RC and Poate JM, J. Appl. Phys. **70**, 3778, 1991.
4. Can N, Townsend PD, Hole DE, Snelling HV, Ballesteros JM and Afonso CN, J. Appl. Phys. **78**, 6737, 1995.
5. Ishikawa Y, Shibata N and Fukatsu S, Appl. Phys. Letts. **68**, 2249, 1996.
6. Polman A, J. Appl. Phys. **82**, 1, 1997.
7. Brongersma ML, Polman A, Min KS and Atawater HA, J. Appl. Phys. **86**, 1999.
8. Priolo F, Franzò G, Pacifici D, Vinciguerra V, Iacona F and Irrera A, J. Appl. Phys. **89**, 264–272, 2001.
9. Hryciw A, Meldrum A, Buchanan KS and White CW, Nucl. Instrum. Methods B **222**, 469, 2004.
10. Nazarov A, Sun JM, Osiyuk IN, Tjagulski IP, Lysenko VS, Skorupa W, Yankopv RA and Gebel T, Appl. Phys. Lett. **86**, 151914, 2005.
11. Wolk JA, Yu KM, Bourret-Courchesne ED and Johnson E, Appl. Phys. Letts. **70**, 2268, 1997.
12. Kanemitsu Y, Tanaka H, Kushida T, Min KS and Atawater HA, J. Appl. Phys. **86**, 1762, 1999.

13. Lee JK, Tewell CR, Schulze RK, Nastasi M, Hamby DW, Lucca DA, Jung HS and Hong KS, Appl. Phys. Lett. **86**, 183111, 2005.
14. Kanemitsu Y and Ishizumi A, J. Luminescence **119/120**, 161, 2006.
15. Amekura H, Umeda N, Sakuma Y, Plaksin OA, Takeda Y, Kishimoto N and Buchal Ch, Appl. Phys. Lett. **88**, 153119, 2006.
16. Buchal Ch, Wang S, Lu F, Carius R and Coffa S, Nucl. Instrum. Methods B **190**, 40, 2002.
17. Lu, F, Carius R, Alam A, Heuken M, Rizzi A, Buchal Ch, Thin Solid Films **425**, 171, 2003.
18. Townsend PD, Chandler PJ and Zhang L, Optical effects of ion implantation, Cambridge University Press, Cambridge, 1994; 2006.
19. Townsend PD and Olivares J, Appl. Surf. Sci. **110**, 275, 1997.
20. Miyakawa M, Kamioka H, Hirano M, Kamiya T, Sushko PV, Shluger A, Matsunami N and Hosono H, Phys. Rev. B **73**, 205108, 2006.
21. Sahoo PK, Gasiorek S and Lieb KP, Appl. Phys. Lett. **87**, 021105, 2005.
22. Keinonen J, Gasiorek S, Sahoo PK, Dhar S and Lieb KP, Appl. Phys. Lett. **88**, 261102, 2006.
23. Lieb KP and Keinonen J, Contemp. Phys. **47**, 305, 2006.
24. Oliver A, Cheang-Wong JC, Crespo A, Hernández JM, Solís C, Muñoz E, Espejel-Morales R, Siejka J, Appl. Phys. Lett. **73**, 1574, 1998.
25. Townsend PD, Chandler PJ, Wood RA, Zhang L, McCallum JC and McHargue CW, Electron. Lett. **26**, 1193, 1990.
26. Chen F, Wang X-L and Wang K-M, Opt. Mater. **29**, 1523, 2007.
27. Feng Chen, J Appl. Physics in Press 2009.
28. Buchal C, Nucl. Instrum. Methods B, **166–167**, 743, 2000.
29. Meldrum A, Boatner LA and White CW, Nucl. Instrum. Methods B **178**, 7, 2001.
30. Townsend PD, Brooks R, Hole DE, Wu Z, Turkler A, Can N, Suarez-Garcia A and Gonzalo J, Appl. Phys. B **73**, 345 2001.
31. Olivares J, Requejo-Isidro JM, del Coso R, de Nalda R, Solis J, Afonso CN, Stepanov AL, Hole DE, Townsend PD and Naudon A, J. Appl. Phys. **90**, 1064, 2001.
32. Hole DE, Townsend PD, Barton JD, Nistor LC and Van Landuyt J, J. Non Cryst. Solids **180**, 266, 1995.
33. Yang B, Townsend PD and Fromknecht R, J. Phys. Condens. Matter **16**, 8377, 2004.
34. Yang B, Townsend PD and Fromknecht R, Nucl. Instrum. Methods B **217**, 60, 2004.
35. Yang B, Townsend PD, Fan Y and Fromknecht R, Nucl. Instrum. Methods B **226**, 549, 2004.

Part V
Equipment and Practice

In this part of the book, the instrumentation and practical techniques needed for nanometre-scale materials characterisation and modification are discussed. This instrumentation has its roots in the tools developed in the 1930s and onwards for fundamental nuclear physics research. The requirement in nanoscience and technology for analysing or processing many different samples has led to a quite different evolutionary path compared to nuclear physics, where many measurements are made on a single specimen. In nanoscience and technology characterization and processing tools, on the other hand, are typically able to handle large numbers of samples and are often instrumented with other complementary surface analysis and processing techniques.

The first step is to produce high quality ion beams suitable for nanoscience research and technology. Different types of accelerators for both modification and analysis of nanostructures are presented in the first chapter. For applications where small beam spots are needed, the ion beam must be focused using magnetic lenses, which are discussed in the next chapter. Ion spectrometers and detectors are another key part of the instrumentation. These are used not only for materials characterisation, but also for end-point and diagnostic sensors for process control. The final chapter discusses readout and control electronics and the new possibilities afforded by modern approaches such as application specific integrated circuits (ASIC) and computer-controlled instrumentation.

Finally, the appendices collect key information about the SI units and a guide to selection of methods.

Micro- and Nanoengineering with Ion Tracks

Christina Trautmann

1 Introduction

In nature, particle tracks are formed over millions of years whenever radioactive elements decay by fission processes, producing energetic fragments of sufficiently large mass ($\sim 100\,\mathrm{u}$) and energy ($\sim 100\,\mathrm{MeV}$).

Fifty years ago, Young recognized the existence of fission tracks in LiF after making damage trails optically visible by chemical etching [1]. Almost simultaneously, Silk and Barnes published the first direct transmission electron micrographs of tracks in mica [2]. In later years, ion tracks have been recorded in many different materials, mainly in insulators such as oxides, spinels, ceramics, and ionic crystals, but also in semiconductors, and in a few selected metals.

For track technology, the capability of forming pores of extremely uniform diameter through chemical etching was most important [3]. This discovery has triggered numerous applications in an extremely broad range of scientific and industrial areas. In geoscience, for instance, the evaluation of the number density of fission tracks accumulated over time in minerals (such as mica, phlogophite, apatite, or zircon) has become a standard dating method for geological samples [4]. Today most applications are based on etched tracks in polymers. Ion track membranes are available as commercial products [5] and are in use, for example, as special filters, cell cultivation substrates, and templates for the synthesis of micro- and nanowires.

The possibility of producing particle tracks under controlled irradiation conditions was recognized at an early stage by means of fission fragments from radioactive sources [3]. Later, the advent of large accelerator facilities provided access to improved irradiation conditions, including ion beams of all elements (from proton up to uranium) and adjustable beam energy. During the last decade, the rapidly increasing activities and recent developments in nanoscience have boosted the interest and created new opportunities for ion track technology [6–9].

C. Trautmann (✉)

GSI Helmholtz Centre for Heavy Ion Research, 64291 Darmstadt, Germany

e-mail: c.trautmann@gsi.de

The superior properties of the second-generation ion track–based nanostructures are closely related to the ability to control and manipulate their number density, length, and diameter. In the energy range of some hundreds of MeV and above, each individual heavy ion induces a straight trail of damage with a width of a few nanometers and a length of typically several tens of micrometers. Tracks are predominantly produced in insulators and consist of material with properties that are drastically changed with respect to the surrounding matrix. Compared to other techniques, track-based nanostructures exhibit an extremely high length-to-diameter ratio of 10^4 and more.

The first part of this chapter characterizes the most important processes accompanying the interaction of energetic heavy ions with matter. Beam parameters and material properties relevant for producing nanostructures are specified. For most applications, ion tracks are processed by chemical etching, in which the track region is preferentially dissolved and transformed into an open channel. The second part describes the state-of-the-art of this technique, encompassing recent developments and experiments. Finally, various applications illustrate how track-etched nanostructures are suitable to exploit novel physical, chemical, and biological properties of materials on the nanometer scale.

2 Track Formation

The slowing down of energetic ion projectiles in matter strongly depends on their velocity. At low kinetic energies, elastic collisions with target atoms prevail, whereas MeV-GeV ions interact mainly with the target electrons. A scheme of the different stopping processes as a function of the kinetic energy is shown in Fig. 1, presenting the energy loss of gold ions in a polycarbonate target.

The mechanisms of damage creation in these two energy regimes are quite different. At low energies, atoms are directly displaced from their lattice sites, provided the energy available is sufficiently high to displace a given atom from its lattice site. Defect creation via elastic collisions occurs in all solids, and follow-up processes such as defect annihilation and aggregation are rather well understood. In contrast, damage creation by swift ions in the electronic energy loss regime is more complex because the atoms are not displaced from their lattice positions in a direct process. At high energies, ion projectiles first transfer their energy to the loosely bound electrons of the solid [10, 11], initiating a cascade of secondary electrons which spreads radially. After some time delay, the energy is gradually transferred into atomic motion, thereby modifying the atomic structure of the solid.

The electronic energy loss has a maximum (Bragg Peak) for kinetic ion energies between 1 and 4 MeV per nucleon (MeV/u). The energy deposition per unit path length increases with mass and can reach values as large as some tens of keV/nm [12]. Via the electron cascade, about half of this energy is deposited within a few nm around the ion path [13], providing in extreme cases sufficient energy to break all atomic bonds in this region.
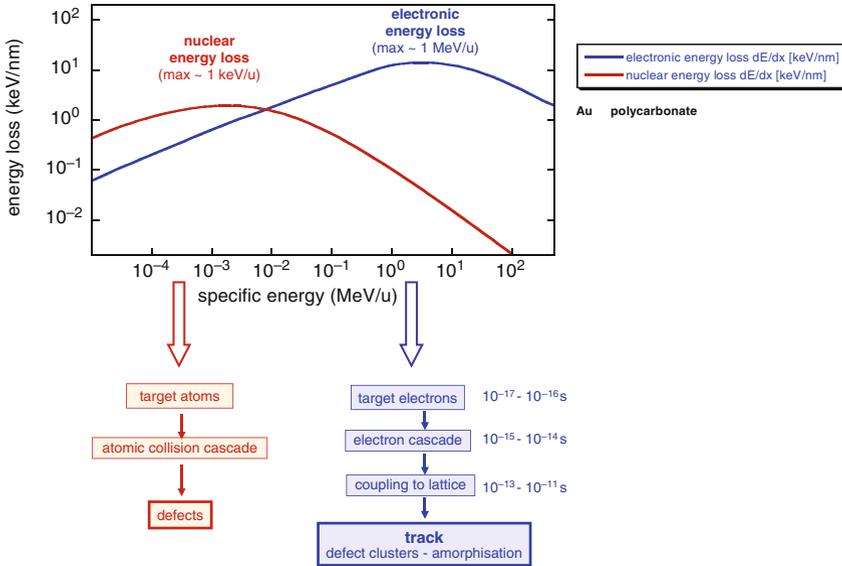
Fig. 1 Energy loss of gold projectiles slowing down in a polycarbonate target estimated with the SRIM-2000 code [12]. For specific energies above 0.1 MeV per nucleon (MeV/u), the interaction is dominated by electronic excitations and ionizations (electronic energy loss); whereas at lower energies, elastic collisions (nuclear energy loss) govern the defect creation process

2.1 Thermal Spike Concept

Although ion track formation has been the subject of research for decades, our current understanding of the complex mechanism is far from being complete. The main unsolved question concerns the spread of the ion-deposited energy as a function of space and time, and its conversion into atomic motion. Besides several other descriptions [14–17], the main theoretical approaches can be grouped into two categories: (1) the coulomb explosion approach, which assumes that coulomb repulsion triggers the movement of ionized atoms [3, 18]; and (2) the thermal spike model, in which track formation is linked to local lattice heating [19–21]. In both models, the key parameters are the number and the mobility of electrons.

In the coulomb explosion, the repulsion between ionized lattice atoms depends on the screening time given by returning electrons, being slow in insulators and rapid in metals. The mechanism was originally proposed by Fleischer et al. [3], when tracks were observed exclusively in insulating materials. Later, when track formation was also evidenced in some metals, several attempts were undertaken to calculate the repulsive forces and estimate if the strength and lifetime are large enough to induce atomic motion in metals [18, 22]. However, to this day, there is no coulomb explosion description available which is sufficiently developed to correctly reproduce experimental observations or provide useful predictions.

In the thermal spike approach, the diffusion of the deposited ion energy is followed in space and time by classical heat transport equations. Estimations of the temperature reached along the ion path depend on details of the coupling between electronic system and lattice and can vary quite noticeably, from tens up to thousands of degrees. A crucial point is the radial spread of the electron cascade dissipating part of the deposited energy into a larger volume and not contributing to local lattice heating.

Fig. 2 Thermal spike model calculations describing the energy dissipation of 200-MeV Au ions in SiO_2 quartz: temperature of electrons (*top*) and atomic subsystem (*bottom*) as a function of time. The different lines represent calculations for four selected cylinder radii around the ion path [20]
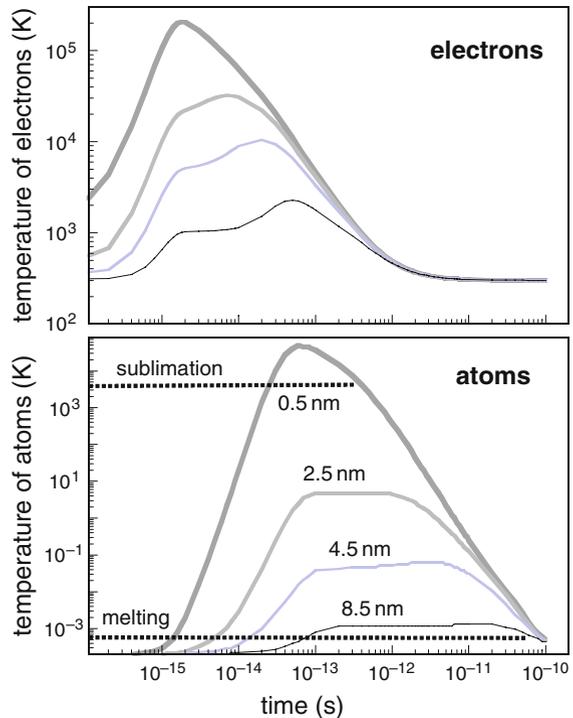
As an illustration, Fig. 2 shows a calculation performed with the inelastic thermal spike model developed by Toulemonde et al. [20]. Because the electronic excitation and atomic motion occur on completely different time scales, the model treats the thermal diffusion equations of both subsystems as quasi-independent, yet coupled via an energy exchange term (electron-phonon coupling). The two graphs illustrate how the energy is first shared and thermalized between the electrons (time scale $10^{-16} - 10^{-13}$ s). With some time delay, the energy is then transferred to the lattice atoms, producing a local thermal spike. For metals and amorphisable insulators, track formation has been linked to the criteria of local melting, i.e., a cylindrical region around the ion trajectory contains sufficient energy to be transformed into the molten phase. During subsequent cooling, this liquid cylinder is rapidly quenched, freezing a highly disordered track region. Track recording materials are characterized by their high efficiency to transfer energy from the electrons to the atoms.

For many materials, this two-temperature model quantitatively describes the evolution of the track radius with electronic stopping and experimental threshold values for the onset of track formation [20]. However, direct confirmation suffers from the fact that most experimental techniques do not have direct access to important parameters such as local electron energy and ionization density. Also, the applicability of the thermal spike concept is sometimes heavily disputed because of the use of equilibrium heat transport due to the lack of information on nonequilibrium thermodynamical material properties [23].

2.2 Material Sensitivity

The formation of tracks typically appears above a critical value of electronic energy loss. This threshold is directly correlated with the radiation sensitivity of a given material, being high for metals and low for insulators, in particular for organic materials such as polymers (Fig. 3).
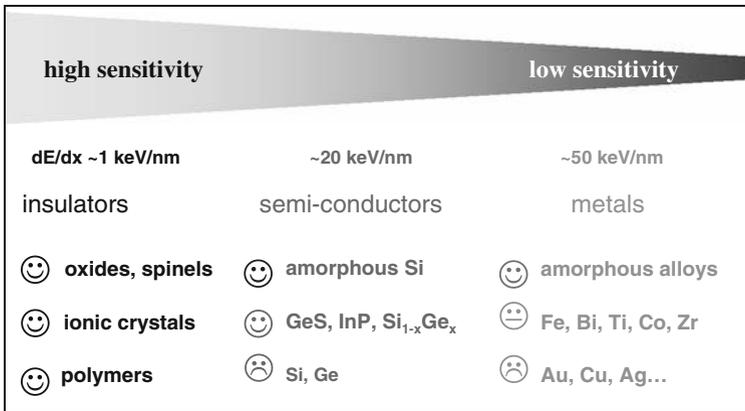
Fig. 3 Track registration sensitivity and range of electronic stopping power threshold for different materials

Tracks have been observed only in a few selected metallic targets (e.g., Bi, Zr, Ti) [21, 24–26] and in various metallic glass alloys [27–29]. Pure semiconductors are insensitive; in Si, e.g., tracks are only formed under the extremely high energy loss of ion clusters [30, 31], but not for monatomic projectiles. Many semiconducting compounds show low susceptibility to track creation [32–35]. Amorphous solids seem to be more sensitive than their crystalline counterparts [36, 37], possibly due to a more efficient energy transfer between the electronic and lattice systems. The best track recorders are insulators, and therefore most ion track studies and application oriented investigations are related particularly to polymers [38–42], but also to oxides [43–47], superconductors [48, 49], spinels [50–52], ceramics [53–55], and ionic crystals (e.g., LiF, CaF_2, NaCl, etc.) [56–60]. Special attention was also given

to layered materials such as mica [32, 61], GeS [32], and highly oriented pyrolytic graphite (HOPG) [62, 63].

2.3 Track Size and Damage Morphology

In many materials, predominantly in insulators, tracks consist of extended amorphous cylinders, as illustrated in Fig. 4. The tracks are embedded in the crystalline matrix, with the interface between the amorphous zone and intact surroundings being rather sharp. Tracks typically have a diameter of several nanometers. Their length is determined by the kinetic energy and the corresponding range of the ions. A rough estimation (neglecting the target mass density) gives a track length of about $10\,\mu m$ per 1 MeV/u kinetic energy of the projectile.

Fig. 4 High-resolution transmission electron microscopy images of single amorphous tracks formed (*left*) in a high-T_c superconductor $Bi_2Sr_2CaCu_2O_8$ (ion passage normal to the layered structure) [49] and (*right*) in the mineral zircon, $ZrSiO_4$ (cross-sectional view) [64]
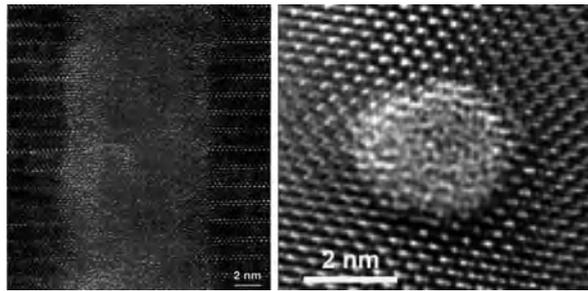
The track region is characterized by a severe modification of the physical and chemical structure (including defects and defect clusters), and by a reduced mass density. In materials sensitive to radiolysis, such as polymers, this effect can be quite significant because of the release of volatile degradation products (e.g., H_2, CO, CO_2, hydrocarbons).

Two of the most important track characteristics are the size and the damage morphology as a function of the electronic energy loss. The most detailed results are available for amorphous tracks in yttrium iron garnet [65] investigated by different techniques for numerous ion species over a large energy loss range. Based also on existing data for other materials [66–68], the following description seems to be of general validity (see also scheme in Fig. 5, left):

- Slightly above the track formation threshold, the damage along the ion trajectory consists of separated extended defects.
- With increasing energy loss, these defect zones percolate into a discontinuous, quasi-cylindrical, and finally continuous but inhomogeneous track.
- At higher energy losses, homogeneous track cylinders are formed. The radius of homogeneous tracks becomes larger when further increasing the energy loss.
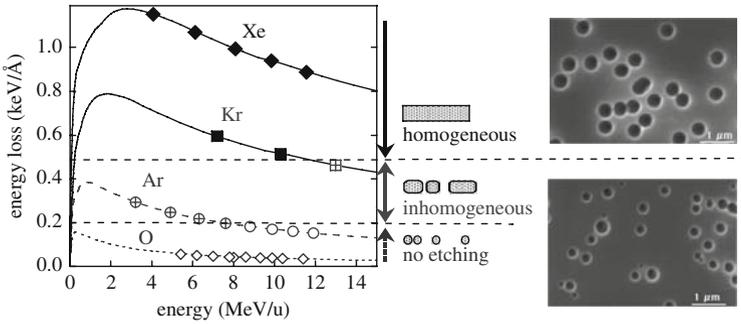
Fig. 5 Energy loss of different ion projectiles in polyimide. The different symbols denote cylindrical (*full*), discontinuous (*crossed*), and spherical (*open*) damage morphology with respective homogenous (*full*), inhomogeneous (*crossed*), and missing (*open*) selective track etching. Etching of homogeneous tracks results in uniform pores (*top* micrograph), whereas tracks with inhomogeneous damage lead to pores of broad size distribution (*bottom* micrograph) [66]

The absolute threshold value is correlated with the sensitivity of the different materials and slightly increases for higher ion velocities [7, 20].

The damage morphology is also important for the production of ion track membranes in which the tracks are preferentially dissolved with a suitable etchant. Monodisperse pores are obtained above a critical threshold where the damage is homogeneous, whereas inhomogeneous damage leads to channels of broad size distribution (see scanning electron microscopy (SEM) image of etched channels in Fig. 5, right) [66].

3 Irradiation Experiments with Swift Heavy Ions

The most exciting aspect of MeV-GeV heavy ions as tools for applications is the fact that each individual projectile produces a cylindrical nanometer-sized track. At high kinetic energy, no straggling occurs and ion tracks can be regarded as well aligned high aspect ratio nanostructures.

Particle beams provided at large accelerator facilities, such as GANIL (Caen, France) or GSI (Darmstadt, Germany), are available in principle for almost all ion species (protons up to uranium). Heavier projectiles are preferred because their high energy loss produces tracks of larger diameter and more pronounced damage. The range of the ions in a specific material is determined by the beam energy. If the range is sufficiently large (e.g., for GeV ions), sample stacks can be irradiated with rather small energy loss variations within the stack.

To expose larger samples, a collimated beam is scanned over several cm^2, or alternatively, the beam is defocused to illuminate a larger sample area. Tracks are usually stochastically distributed (Fig. 6, left). The ion fluence to be applied can range from a single ion impact per sample (Fig. 6, right) up to heavily overlapping tracks (above $\sim 10^{12}$ ions/cm^2).
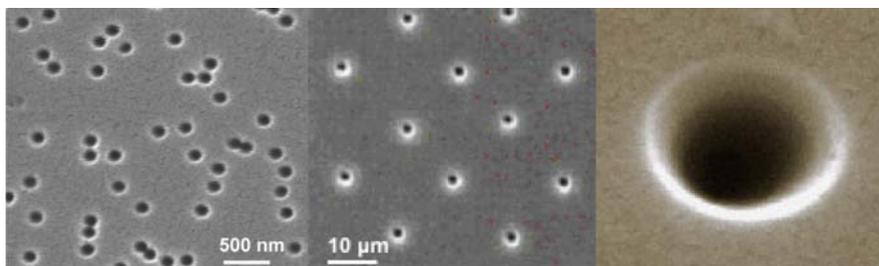
Fig. 6 Accelerated ions are randomly distributed over the selected target area (*left*), but they can also be precisely placed by using a microprobe [71] (*center*). The beam flux and fluence to be applied can be adjusted over a wide range, including the controlled exposure to one single ion per sample (*right*). All micrographs show ion tracks enlarged to pores by chemical etching

An essential part of the irradiation experiments is the adjustment and monitoring of the beam flux. In order to avoid target heating, in particular for rather sensitive materials such as polymers, the flux of heavy GeV ions (e.g., Au, Pb, or U) should be limited to 1–2×10^8 ions/cm^{2}s. For light ions ($< Ar$), this value can be one order of magnitude larger.

For nondestructive online control of high energy beams (above $\sim 1\,MeV/u$), a most suitable system is a SEcondary Electron TRAnsmission Monitor (SEE-TRAM) [69]. It consists of three thin metallic foils placed in front of the sample. The total thickness of the foils is such that the detector is transparent for the ions, and the energy loss is small (e.g., in three 1-μm thick Al foils, a 10 MeV/u beam loses about 0.3 MeV/u). When passing through the foils, the beam produces electrons which are collected. This electron current serves as a relative online signal but needs to be calibrated via absolute ion beam measurements with a Faraday cup.

For low fluences ($< 10^4$ ions/cm^2) or single ion irradiations, the sensitivity of the SEETRAM is insufficient. Instead, a solid-state particle detector (e.g., a silicon surface barrier detector) is placed behind the sample detecting the passage of individual ions. To produce single-tracked membranes (Fig. 6, right), the sample is irradiated through a small circular aperture (diameter, e.g., 200 μm). The beam is strongly defocused and adjusted in such a way that single ions pass through the aperture with a frequency of about 1 Hz. Such a setting allows one to quickly switch off the beam by an electrostatic chopper system as soon as the detector has registered a single ion impact.

Individual ions can also be placed on predefined positions using a heavy-ion microprobe [70]. The facility in operation at GSI (Darmstadt, Germany) allows positioning with an accuracy of ca. 1 μm (Fig. 6, center). The applications of microprobe irradiations include writing specific patterns [71], testing microelectronic circuits [70], and, more recently, delivering a preset number of ions to the nuclei of individual living cells [72].

4 Chemical Track Etching

By far the most frequent application of ion tracks is linked to chemical etching, which is based on the preferential dissolution of the damaged track material. In an appropriate chemical solution, the etch rate along the ion track can be orders of magnitude larger than the bulk etch rate of the surrounding matrix material. For a high track-to-bulk etch rate ratio, the geometry of the resulting pore is essentially cylindrical, whereas a low ratio leads to channels of conical (single-side etching) or hourglass (double-side etching) shape. The diameter and geometry of the pores are controlled by the etching time and by specific details of material and etchant (e.g., concentration, temperature, pH, additives, applied voltage).

For polymers, the exposure to UV light has been shown to significantly improve the track etchability and pore homogeneity [73].

5 Track-Etched Membranes

The major applications of ion track technology are based on nanopore fabrication in commercial polymer foils (e.g., polycarbonate, polyethylene terephthalate, polyimide, etc.); other materials, such as mica and vitreous or crystalline SiO_2 quartz [74], are also suitable. Although Si can not be structured by ion tracks directly, a thermally grown oxide layer or a thin polymer layer spin-coated on a Si substrate make track structuring compatible with the silicon technology [75].

The specific advantages are linked to the excellent control of the membrane porosity determined by the number of tracks per unit area (irradiation fluence) and the diameter of the etched pores. To date, commercial ion track membranes are available with pore diameters between 15 nm and several μm and with a wide range of pore density [76, 77]. Single-pore membranes with a pore aperture of only a few nanometers are feasible, as well as membranes of porosity as large as about 20%. This upper limit is given by the mechanical stability of the membrane when pore overlapping becomes dominant. Polymer membranes offer the additional advantage that the wall of the pores can be modified e.g. by grafting [78] or by chemically changing the hydrophobicity.

Most of the applications are linked either to directly perforating material or to depositing material into the pores of etched ion track membranes. To give a short glimpse of application-oriented activities in this field, the following two sections describe some of the possibilities recently developed.

5.1 Integrating Tracks into Microsystems

Porous materials for MEMS (Micro-Electro-Mechanical Systems) structures have gained interest in the field of filtration and drug delivery. Combining polymer-based microfabrication with ion track technology allows the direct integration of pores in a flexible microdevice (Fig. 7). By adjusting the energy of the ion beam, the
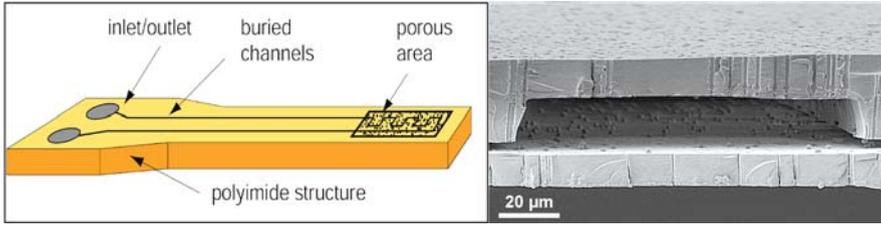
Fig. 7 Microfluidic system (*left*) with an integrated porous membrane area (*right*) suitable for cross-flow filtration of particles and molecules in fluids or as bioimplants for drug delivery [79]

ions traverse either only the top layer or both the top and the bottom layer of the device. By using an irradiation mask, which is sufficiently thick to stop the ions, selected filter zones can be produced. The polyimide microfluidic channel structure is produced by a spin-coating and lamination technique. The upper and lower layers of the channel are perforated by ion irradiation and subsequent track etching [79].
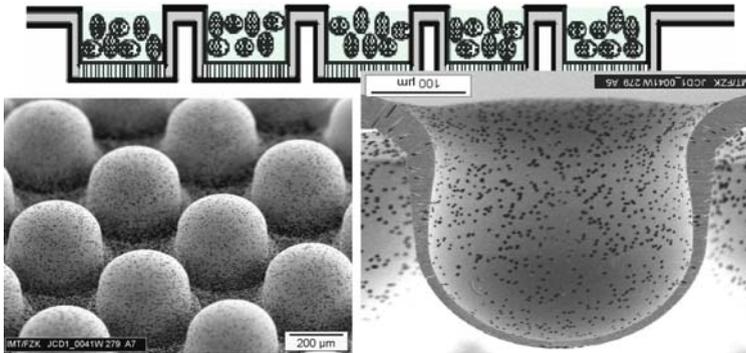
Fig. 8 Scheme of microcontainer set for cell cultivation (*top*). The device is produced by thermoforming of ion irradiated polymer sheets and subsequent track etching, creating cylindrical pores (*black dots* in both micrographs). The cross-section of an inverted single compartment (*bottom right*) shows the perforated walls, which provide improved supply and flushing with liquid nutrients [80]

Figure 8 presents another example of a microdevice structured by means of etched ion tracks. The optically transparent scaffold for the growth of biocells was provided with nanopores for sufficient nutrient media supply [80]. In analogy to macroscopic thermoforming of pretextured or printed semifinished products, these microcontainers are fabricated by thermoforming a thin polycarbonate film preirradiated with energetic ions. The ion irradiation before the thermoform process has the particular advantage that flat, unstructured polymer sheets may be ion exposed over a large area and at the requested parameter set (range, fluence) before they are subjected to three-dimensional molding. The forming process takes place in the viscoelastic state of the polymer, at a temperature low enough to avoid erasing the tracks by thermal annealing.

Integration of nanopores in MEMS systems is still in an early stage, but it is clear that the ion track technology offers promising perspectives to fabricate more complex devices with novel capabilities, including nanofluidic systems, nanoreactors, and extremely small and compact analysis systems.

5.2 Single-Pore Membrane

There is increasing research activity using artificial nanopores for transport measurements of ions and molecules. The interest is linked to the fact that living cells use protein nanopores and channels embedded within the cell membrane to communicate with the extracellular environment. Such protein channels have a length and a diameter of a few nanometers; they open and close in response to electrical, chemical, and other kinds of stimulus; and they are essential for almost all physiological processes of living organisms.

A number of limitations have restricted experimental investigations with natural membranes. They are mainly associated with the short life time of the extremely fragile lipid bilayers and their sensitivity to environmental conditions. To explore more details of ion transport mechanisms, more robust test systems with artificial nanopores are therefore of large interest. At present, different solid-state systems are under study, including nanopores manufactured by lithography, drilling holes with a focused ion beam, and electron beam sculpting [81]. Also, ion track technology offers very suitable conditions for fabricating membranes with single nanochannels of increased durability and control over pore size and geometry.

Besides serving as mimics of biological channels, single nanopores are also very promising in biotechnology, because many separation and sensing processes require pore openings of a few nanometers, the typical size of biomolecules.

In the recent past, the fabrication of single-nanopore membranes by means of ion track technology has benefited from major advances related to: (1) tracking of polymer films with a single particle, (2) conductivity-controlled etching in combination with chemical stopping of the etch process [82, 83], and (3) etching conditions allowing conical pore geometries.

For the online controlled etching procedure, the ion irradiated sample is clamped between two halves of a conductivity cell (Fig. 9) [82]. One half cell is filled with the etchant, while the other contains a chemical stopping media which neutralizes the etchant upon pore breakthrough and thus slows down further etching significantly. To provide additional control of the nascent pore, a voltage is applied to the electrodes inserted into each compartment, with the polarity being such that the anode is in the etch solution. Once the pore opens, the electric field drives the etchant ions away from the pore tip. Measuring the ionic current flowing through the pore by a high quality picoamperemeter allows the controlled synthesis of asymmetric pores with tips of extremely small diameters, down to a few nanometers.

The sensing and transport properties of conical nanopores strongly depend on size and geometry (cone angle, membrane thickness). Characterization is therefore
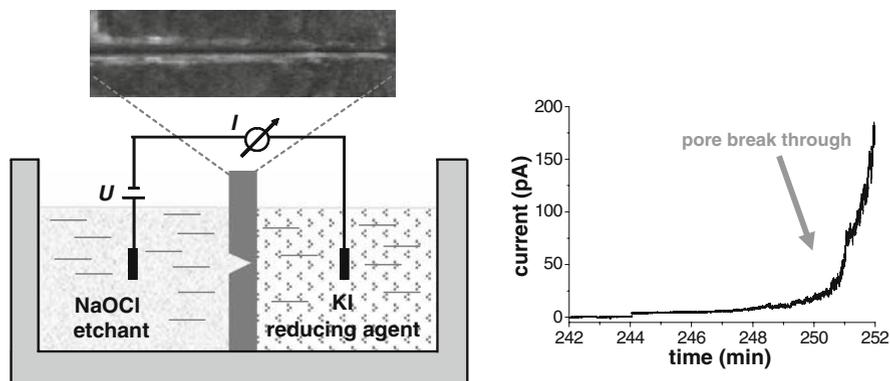
Fig. 9 Scheme of conductometric etching cell and scanning electron micrographs of conical single nanopore (larger pore opening at *left*) produced by exposing one side of the ion track to a suitable etchant (e.g., for polyimide: NaOCl) [82]. Pore breakthrough is identified by a steep increase in conductance (*right*). At that moment, the reducing agent (e.g., KI) neutralizes the etchant (OCl^- + $2H^+ + 2I^- \rightarrow I_2 + Cl^- + H_2O$) and slows down further pore growth

important, but presents its own challenges because of the small scale involved. Conductance measurements possible by filling the pore with an aqueous electrolyte (e.g., KCl) allow indirect estimations of the effective pore size. There is, however, a caveat, because the relationship between the conductivity of the electrolyte in a nanopore and the conductivity of the bulk electrolyte is complicated and should consider effects such as surface charges, inhomogeneous ion concentration, and the electrical double layer at the pore wall. In addition, details of the pore geometry are not well known and are only accessible by independent measurements of the bulk and track etch rate.

The conductometry cell is not only useful for real-time monitoring and stopping of the etching process, but is also suitable to investigate specific electrical phenomena. When a potential difference is applied across the membrane, the ionic flow through the electrolyte-filled nanopore generates a measurable electric current signal. The response of ion current to an analyte of interest or to different pore properties (size, charge, etc.) can be measured as a function of potential (current–voltage curves) or as a function of time (current–time transients).

Nanopores have been demonstrated to act as ion current rectifiers, exhibiting nonlinear current-voltage curves. Excess surface charges at the small tip modify the electrical field and lead to preferential ionic flow from one side of the pore compared to the reverse direction [84].

Another consequence of the conical pore geometry is that the resistance focuses on the region of smallest aperture within about 100 nm effective length. The ion current is thus extremely sensitive to analyte species present in or near the nanopore tip [85].

The surface of pores in polymer membranes can easily be modified because the track etching process leaves dangling endgroups that are charged and can be

functionalized [86]. For sensor applications, e.g., the pore specificity can be tailored by attaching certain molecules that can interact with the desired analyte. Surface modifications also offer the chance to fine-tune the pore size by gold plating through electroless deposition and to make thiol chemistry accessible, which permits chemisorption of functionalized species [87].

Single conical nanopores serve as resistive-pulse sensors detecting the translocation of a small molecule as a current-blocking event. Ionic current blockages measured in voltage biased nanofabricated pores have been used to sense nanoparticles and molecules such as proteins or DNA [88, 89]. There has been remarkable progress; however, the work on nanopore-based DNA analysis and surface modification and functionalization is only at the beginning.

6 Templates for Nanowire Fabrication

Ion beam technology provides a successful route for growing nanostructures by filling material into track-etched channels. The replication of etched pores benefits from the same advantages as the template itself, i.e., adjustable areal density, pore monodispersity, shape tuning, excellent control of diameters and lengths, and thus access to a wide range of aspect ratios. As templates, mainly polymer membranes are utilized. For pore filling, very different materials are possible, including metals, semiconductors, and insulators [76]. Electroless deposition plates only the pore walls producing tubes. Multilayered wires are formed by applying alternating voltages forming, e.g., layers of ferro- and diamagnetic metals (e.g., Co/Cu), which are extremely interesting with respect to their giant-magneto-resistive behavior [90]. Also nonmetallic materials, such as conductive polymers [91], and even alkali halide [92] or organic crystals [93], have been synthesized.
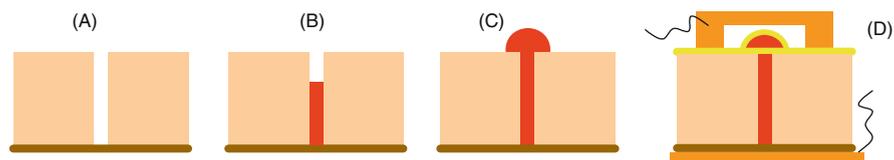
Fig. 10 Synthesis of metal nanowires by the template technique: (**A**) etching of ion tracks, (**B**) deposition of metallic surface cathode and electrochemical filling of metallic material into the pores. The wire length can be tuned by terminating the deposition process at a preset time. Dissolving the polymer matrix leads to freestanding needles (see Fig. 11). (**C**) Under prolonged metal deposition, caps are formed on the upper surface, suitable as contact buds, e.g., for electrical measurements on single wires (**D**)

For the fabrication of metal nanowires (e.g., Cu, Au, Pt, Ni, Co, Fe, Bi), electrochemical deposition is the most suitable process. It consists of several steps (Fig. 10): First, ion tracks in a given membrane are chemically etched and enlarged to pores of a certain shape and size; then a thin metallic layer is sputtered onto one side of the polymer membrane, reinforced galvanostatically with copper. This metal

layer provides a stable substrate for the needle growth and serves at the same time as a cathode. The template is immersed in a galvanic bath and a constant or alternating voltage is applied. Monitoring the electrical current allows online control of the wire growth. Complete filling of the pore is characterized by a rapid increase of the deposition current due to cap formation at the wire tip on the template surface. The wire length can be adjusted by terminating the deposition process at a preset time or integrated charge. For further use, the nanostructures can remain inside the pores or they can be released from the template and harvested by dissolving the polymer in an organic solvent (Fig. 11).

Fig. 11 Ensembles of freestanding metal microwires (with and without caps) obtained after dissolving the polymer template [94]
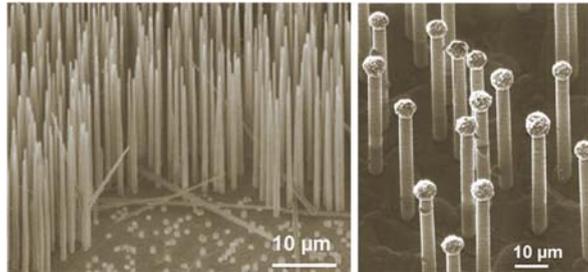
By a proper choice of the synthesis parameters (composition and concentration of electrolyte, voltage applied, and temperature), the grain size [95] and the crystallographic orientation of the nanowires can be adjusted [96].

The enormous interest in nanowires is based on a variety of promising applications. Examples include interconnects for nanoelectronics (Fig. 12) [97], magnetic and thermoelectric devices [98], chemical or biological sensors, and emitters in field emission arrays [99, 100]. At the same time, they are considered to be excellent objects to study fundamental physical phenomena such as electron transport [101, 102], magnetoresistance [103], optical absorption [104], or quantum-size effects [105].
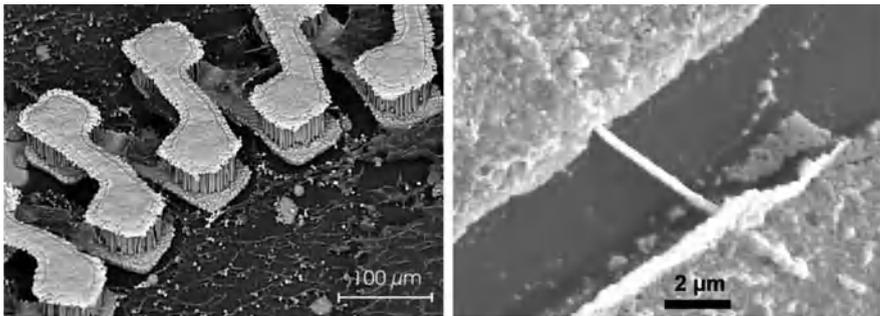
Fig. 12 3-D nanowire circuits (*left*) [97] and nanowire bridge (*right*) fabricated by means of template method and conventional lithographic contacts [106]

The technical application of smaller and smaller wires requires a thorough understanding of size-dependent phenomena. For example, before implementing nanowires in nanoelectronic and sensor devices, the response of the nanomaterial during thermal annealing should be known. A study of copper and gold nanowires showed, for instance, that their thermal stability depends on the wire diameter; and that fragmentation into small segments (Rayleigh instability), and finally into a chain of spherical droplets, occurs far below the melting temperature of bulk material [107].

Also, the electrical properties of metal nanowires are being intensively investigated to understand how the properties of solids deviate from macroscopic behavior when their dimensions are comparable to characteristic length scales such as the electron mean free path. The main experimental challenge on the way towards electrical characterization is the prevention of oxidation [108] and the fabrication of reliably stable electrical contacts. Here, ion track technology offers an interesting solution by using single nanowires embedded in the template. As illustrated in Fig. 11 (scheme D), the wire is encapsulated, and thus protected by the polymer matrix, and is contacted by planar electrodes sputtered onto the template surface. Based on this approach, electronic quantum effects as well as the influence of grain size and wire diameter have been investigated in a broad temperature regime [102, 109].

7 Summary

The past two decades have witnessed rapid advances in the ability to control the structure of matter at the nanoscale. Ion track technology offers the potential to exploit novel physical, chemical, and biological properties of materials at this small scale. It also opens the door to a wide gamut of applications in various fields, such as sensor technology, molecular biology, or biomedicine. The implementation of nanopores in microsystems is still in an early stage, but promising perspectives for more complex devices with novel capabilities, including nanofluidic systems, nanoreactors, and extremely small and compact analysis systems, have been demonstrated.

The advantages and superior properties of ion track–based micro- and nanostructures are related to their large length-to-diameter ratio and to the ability to control and manipulate their number density, shape, length, and diameter.

However, there are various difficulties to overcome before ion beams are fully accepted as a versatile tool, in particular for industrial applications. One obstacle is certainly that ion beams of several MeV per nucleon require large accelerator facilities with limited beam time. Apart from the fact that the access problem might possibly be solved, it should be emphasized that smaller, more prevalent accelerators, delivering ions of some tens of MeV and ion masses around 100, are interesting for applications. They are suitable in particular for those projects in which small ion ranges of only several μm are sufficient [7].

Finally, it should be mentioned that there is still a substantial need for a better understanding of the physical processes involved in track formation. From the experimental side, techniques need to be developed to access short-time phenomena, e.g., by sophisticated online spectroscopy. From the theoretical side, it would be extremely important to extend microscopic and macroscopic models, e.g., by including hydrodynamic processes such as pressure spikes, as well as recrystallization occurring at the interface between track and matrix. Molecular dynamics calculations should be performed for materials under conditions closer to the experimental situation. It would be a significant improvement if the internal structure and binding character of different target materials such as metals, oxides, or ionic crystals could be taken into account. Future simulations also should try to unveil the limitations of continuum mechanics, but they need to deal with a great number of atoms, electronic excitations, and the concomitant variations in the interatomic potentials.

References

1. Young DA (1958) Nature 182:375–377
2. Silk ECH, Barnes RS (1959) Phil. Mag. 4:970–972
3. Fleischer RL, Price PB, Walker RM (1975) Nuclear Tracks in Solids. Principles and Applications. University of California Press, Berkeley
4. Wagner GA, Van den haute P (1992) Fission Track-Dating. Kluwer Academic Publishers, Dordrecht
5. e.g., Oxyphen GmbH (www.oxyphen.com/), Whatman, GE Healthcare (www.whatman.com), Osmonics, General Electric Company (www.osmolabstore.com/), it4ip company (www.it4ip.be/index.htm)
6. Spohr R (1990) Ion Tracks and Microtechnology, Principles and Applications. Vieweg Verlag, Braunschweig.
7. Toulemonde M, Trautmann C, Balanzat E, Hjort K, Weidinger A (2004) Nucl. Instrum. Methods B 216:1–8
8. Spohr R (2005) Radiat. Measure. 40:191–202
9. Apel P (2003) Nucl. Instrum. Methods B 208:11–20
10. Schiwietz G, Luderer E, Xiao G, Grande PL (2001) Nucl. Instrum. Methods B 175–177:1–11
11. Sigmund P (2006) Stopping of Swift Iones: Solved and Unsolved Problems in: Ion Beam Science, Solved and Unsolved Problems. The Royal Danish Academy of Sciences and Letters, Copenhagen, Mat. Fys. Med. 52–2:557–594
12. TRIM & SRIM code, Ziegler JF, Biersack JP, Littmark U (1985) The Stopping and Ranges of Ions in Solids. Pergamon Press, New York (http://www.srim.org/)
13. Gervais B, Bouffard S (1994) Nucl. Instrum. Methods B 88:355–364
14. Watson CC, Tombrello TA (1985) Radiat. Eff. 89:263–283
15. Stampfli P (1996) Nucl. Instrum. Methods B 107:138–145
16. Itoh N (1996) Nucl. Instrum. Methods B 116:33–36
17. Schiwietz G, Luderer E, Grande PL (2001) Appl. Surf. Sci. 182:286–292
18. Lesueur D, Dunlop A (1993) Radiat. Eff. Defects Solids 126:163–172

19. Toulemonde M, Dufour Ch, Meftah A, Paumier E (2000) Nucl. Instrum. Methods B 166:903–912
20. Toulemonde M, Assmann W, Dufour C, Meftah A, Studer F, Trautmann C (2006) Experimental phenomena and thermal spike model description of ion tracks in amorphisable inorganic insulators. In: Ion Beam Science, Solved and Unsolved Problems. The Royal Danish Academy of Sciences and Letters, Copenhagen, Mat. Fys. Med. 52–2:263–292
21. Wang ZG, Dufour C, Paumier E, Toulemonde M (1994) J. Phys. Condens. Matter 6: 6733–6750
22. Klaumünzer S, Hou M-d, Schumacher G (1986) Phys. Rev. Lett. 57:850–853
23. Klaumünzer S (2006) Thermal spike models for ion track physics: a critical examination. In: Ion Beam Science, Solved and Unsolved Problems. The Royal Danish Academy of Sciences and Letters, Copenhagen, Mat. Fys. Med. 52–2:293–328
24. Barbu A, Dunlop A, Lesueur D, Averback RS (1991) Europhys. Lett. 15:37–42
25. Dunlop A, Dammak H, Lesueur D (1996) Nucl. Instrum. Methods B 112:23–25
26. Iwase A, Ishino S (2000) J. Nucl. Mater. 276:178–185
27. Hou M-D, Klaumünzer S, Schumacher G (1990) Phys. Rev. B 41:1144–1157
28. Trautmann C, Andler S, Brüchle W, Toulemonde M, Spohr R (1993) Radiat. Eff. Defects Solids 126:207–210
29. Rizza G, Dunlop A, Jaskierowicz G, Kopcewicz M (2004) J. Phys. Condens. Matter 16:1547–1561
30. Dunlop A, Jaskierowicz G, Della-Negra S (1998) Nucl. Instrum. Methods B 146:302–308
31. Canut B, Bonardi N, Ramos SMM, Della-Negra S (1998) Nucl. Instrum. Methods B 146:296–301
32. Vetter J, Scholz R, Dobrev D, Nistor L (1998) Nucl. Instrum. Methods B 141:747–752
33. Kucheyev SO, Timmers H, Zou J, Williams JS, Jacadish J, Li G (2004) J. Appl. Phys. 95:5360–5365
34. Wesch W, Kamarou A, Wendler E, Undisz A, Rettenmayr M (2007) Nucl. Instrum. Methods B 257:283–286
35. Mansouri S, Marie P, Dufour C, Nouet G, Monnet I, Lebius H (2008) Nucl. Instrum. Methods B 266:2814–2818
36. Hedler A, Klaumünzer S, Wesch W (2006) Nucl Instrum. Methods B 242:85–87
37. Klaumünzer S, Li CL, Löffler S, Rammensee M, Schumacher G, Neitzert HCh (1989) Radiar. Eff. Defects Solids 108:131–135
38. Steckenreiter T, Balanzat E, Fuess H, Trautmann C (1999) Nucl. Instrum. Methods B 151:161–168
39. Apel PY, Blonskaya IV, Oganessian VR, Orelovitch O, Trautmann C (2001) Nucl. Instrum. Methods B 185:216–221
40. Papaléo RM, Leal R, Trautmann C, Bringa EM (2003) Nucl. Instrum. Methods B 206:7–12
41. Dehaye F, Balanza, E, Ferain E, Legras R (2003) Nucl. Instrum. Methods B 209:103–112
42. Adla A, Fuess H, Trautmann C (2003) J. Polym. Sci. Part B Polym. Phys. 41:2892–2901
43. Toulemonde M, Fuchs G, Nguyen N, Studer F, Groult D (1987) Phys. Rev. B 35:6560–6569
44. Wiss T, Matzke Hj, Trautmann C, Toulemonde M, Klaumünzer S (1997) Nucl. Instrum. Methods B 122:583–588
45. Benyagoub A (2006) Nucl. Instrum. Methods B 245:225–230
46. Jensen J, Razpet A, Skupiński M, Possnert G (2006) Nucl. Instrum. Methods B 245:269–273
47. Lang M, Zhang F, Lian J, Trautmann C, Neumann R, Ewing RC (2008) Earth Planet. Sci. Lett. 269:291–295
48. Zhu Y, Cai ZX, Budhani RC, Suenaga M, Welch DO (1993) Phys. Rev. B 48:6436–6450
49. Wiesner J, Træholt C, Wen J-G, Zandbergen H-W, Wirth G, Fuess H (1996) Physica C Superconduct. Appl. 268:161–172
50. Zinkle SJ, Skuratov VA, Hoelzer DT (1999) Mater. Res. Soc. Symp. Proc. 540:299–304
51. Thomé L, Jagielski J, Gentils A, Nowicki L, Garrido F (2006) Nucl. Instrum. Methods B 242:643–645

52. Yasuda K, Yamamoto T, Seki S, Shiiyama K, Matsumura S (2008) Nucl. Instrum. Methods B 266:2834–2841
53. Weber WJ, Ewing RC, Catlow CRA, Diaz De La Rubia T, Hobbs LW, Kinoshita C, Matzke Hj, Motta AT, Nastasi M, Salje EKH, Vance ER, Zinkle SJ (1998) J. Mater. Res. 13: 1434–1484
54. Zinkle SJ, Skuratov VA, Hoelzer DT (2002) Nucl. Instrum. Methods B 191:758–766
55. Skuratov VA, Zinkle SJ, Efimov AE, Havancsak K (2003) Nucl. Instrum. Methods B 203:136–140
56. Trautmann C, Boccanfuso M, Benyagoub A, Klaumünzer S, Schwartz K, Toulemonde M (2002) Nucl. Instrum. Methods B 191:144–148
57. Schwartz K, Trautmann C, El-Said AS, Neumann R, Toulemonde M, Knolle W (2004) Phys. Rev. B 70:184104.1–184104.8
58. El-Said AS, Cranney M, Ishikawa N, Iwase A, Neumann R, Schwartz K, Toulemonde M, Trautmann C (2004) Nucl. Instrum. Methods B 218:492–497
59. Enculescu M, Schwartz K, Trautmann C, Toulemonde M (2005) Nucl. Instrum. Methods B 229:397–405
60. Khalfaoui N, Rotaru CC, Bouffard S, Toulemonde M, Stoquert JP, Haas F, Trautmann C, Jensen J, Dunlop A (2005) Nucl. Instrum. Methods B 240:819–828
61. Albrecht D, Armbruster P, Spohr R, Roth M, Schaupert K, Stuhrmann H (1985) Appl. Phys. A Solids Surf. 37:37–46
62. Liu J, Neumann R, Trautmann C, Müller C (2001) Phys. Rev. B 64:1841151–1841157
63. Dunlop A, Jaskierowicz G, Ossi PM, Della-Negra S (2007) Phys. Rev. B 76:155403/1–155403/14
64. Lang M, Lian J, Zhang F, Hendriks BWH, Trautmann C, Neumann R, Ewing RC, Earth and Planetary Science Letters 274 (2008) 355–358
65. Meftah A, Brisard F, Costantini JM, Hage-Ali M, Stoquert JP, Studer F, Toulemonde M (1993) Phys. Rev. B 48:920–925
66. Trautmann C, Bouffard S, Spohr R (1996) Nucl. Instrum. Methods B 116:429–433
67. Trautmann C, Dufour C, Paumier E, Spohr R, Toulemonde M (1996) Nucl. Instrum. Methods B 107:397–402
68. Jensen J, Razpet A, Skupiński M, Possnert G (2006) Nucl. Instrum. Methods B 245:269–273
69. Jurado B, Schmidt K-H, Behr K-H (2002) Nucl. Instrum. Methods A:483 603–610
70. Fischer BE, Heiss M, Cholewa M (2003) Nucl. Instrum. Methods B 210:285–291
71. Fischer BE, Heiß M (2007) Nucl. Instrum. Methods B 260:442–444
72. Heiß M, Fischer BE, Jakob B, Fournier C, Becker G, Taucher-Scholz G (2006) Radiat. Res. 165:231–239
73. Pépy G, Boesecke P, Kuklin A, Manceau E, Schiedt B, Siwy Z, Toulemonde M, Trautmann C (2007) Journal of Applied Crystallography 40:s388–s392
74. Musket RG, Yoshiyama JM, Contolini RJ, Porter JD (2002) Journal of Applied Physics 91:5760–5764
75. Skupiński M, Jensen J, Johansson A, Possnert G, Boman M, Hjort K, Razpet A (2007) J. Vacuum Sci. Technol. B Microelectron. Nanometer Struct. 25:862–867
76. Dauginet-De Pra L, Ferain E, Legras R, Demoustier-Champagne S (2002) Nucl. Instrum. Methods B 196:81–88
77. Sirijarukul S, Balanzat E, Vasina EN, Déjardin P (2007) J. Membrane Sci. 296:185–194
78. Cuscito O, Clochard M-C, Esnouf S, Betz N, Lairez D (2007) Nucl. Instrum. Methods B 265:309–313
79. Metz S, Trautmann C, Bertsch A, Renaud Ph (2004) J. Micromech. Microeng. 14:324–331
80. Giselbrecht S, Gietzelt T, Gottwald E, Trautmann C, Truckenmüller R, Weibe-zahn KF, Welle A (2006) Biomed. Microdev. 8:191–199
81. Healy K, Schiedt B, Morrison IP (2007) Nanomedicine 2:875–897
82. Siwy Z, Apel P, Baur D, Dobrev DD, Korchev YE, Neumann R, Spohr R, Trautmann C, Voss K-O (2003) Surf. Sci. 532–535:1061–1066

83. Apel PYu, Korchev YuE, Siwy Z, Spohr R, Yoshida M (2001) Nucl. Instrum. Methods B 184:337–346
84. Siwy Z, Fuliński A (2004) Am. J. Phys. 72:567–574
85. Harrell CC, Siwy ZS, Martin CR (2006) Small 2:194–198
86. Ali M, Schiedt B, Healy K, Neumann R, Ensinger W (2008) Nanotechnology 19:085713/1–6
87. Siwy Z, Trofin L, Kohli P, Baker LA, Trautmann C, Martin CR (2005) J. Am. Chem. Soc. 127:5000–5001
88. Mara A, Siwy Z, Trautmann C, Wan J, Kamme F (2004) Nano Lett. 4:497–501
89. Martin CR, Siwy ZS (2007) Science 317:331–332
90. Gravier L, Wegrowe J-E, Wade T, Fabian A, Ansermet J-P (2002) IEEE Trans. Magnet. 38:2700–2702
91. Duvail J-L, Dubois S, Demoustier-Champagne S, Long Y, Piraux L (2008) Int. J. Nanotechnol. 5:838–850
92. Enculescu M, Enculescu I, Sima M, Neumann R, Trautmann C (2007 J. Optoelectron. Adv. Mater. 9:1561–1563
93. Dobrev D, Baur D, Neumann R (2005) Appl. Phys. A 80:451–456
94. Dobrev D, Vetter J, Neumann R, Angert N (2001) J. Vac. Sci. Technol. B 19:1385–1387
95. Toimil Molares ME, Buschmann V, Dobrev D, Neumann R, Scholz R, Schuchert IU, Vetter J (2001) Adv. Mater. 13:62–65
96. Maurer F, Brötz J, Karim S, Toimil Molares ME, Trautmann C, Fuess H (2007) Nanotechnology 18:135709
97. Lindeberg M, Yousef H, Öjefors E, Rydberg A, Hjort K (2004) Mater. Res. Soc. Symp. Proc. 833:255–264
98. Koukharenko E, Li X, Nandhakumar I, Schiedt B, Trautmann C, Speed J, Tudor MJ, Beeby SP, White NM (2008) Electronics Lett. 44:500–501
99. Xavier S, Mátéfi-Tempfli S, Ferain E, Purcell S, Enouz-Védrenne S, Gangloff L, Minoux E, Hudanski L, Vincent P, Schnell J-P, Pribat D, Piraux L, Legagneux P (2008) Nanotechnology 19:215601
100. Dangwal A, Pandey CS, Müller G, Karim S, Cornelius TW, Trautmann C (2008) Appl. Phys. Lett. 92:063115/1–3
101. Cornelius TW, Picht O, Müller S, Neumann R, Völklein F, Karim S, Duan JL (2008) J. Appl. Phys. 103:103713/1–5
102. Cornelius TW, Toimil-Molares ME, Karim S, Neumann R (2008) Phys. Rev. B 77:125425/1–6
103. Gravier L, Serrano-Guisan S, Reuse F, Ansermet J-P (2006) Phys. Rev. B 73:052410/1–11
104. Neubrech F, Kolb T, Lovrincic R, Fahsold G, Pucci A, Aizpurua J, Cornelius TW, Toimil-Molares ME, Neumann R, Karim S (2006) Appl. Phys. Lett. 89:253104/1–3
105. Cornelius TW, Toimil-Molares ME, Neumann R, Fahsold G, Lovrincic R, Pucci A, Karim S (2006) Appl. Phys. Lett. 88:103114/1–3
106. Cornelius TW, Völklein F, Toimil-Molares ME, Karim S, Neumann R (2006) Proceedings of European Conferece on Thermoelectrics, Cardiff.
107. Karim S, Toimil-Molares ME, Ensinger W, Balogh AG, Cornelius TW, Khan EU, Neumann R (2007) J. Phys. D Appl. Phys. 40:3767–3770
108. Toimil Molares ME, Höhberger EM, Schaeflein Ch, Blick RH, Neumann R, Trautmann C (2003) Appl. Phys. Lett. 82:2139–2141
109. Cornelius TW, Toimil-Molares ME, Neumann R, Karim S (2006) J. Appl. Phys. 100:114307

Ion Accelerators for Nanoscience

Ragnar Hellborg and David Hole

1 Introduction and History

Accelerators can be classified into different principle designs such as electrostatic accelerators, cyclotrons, synchrotrons, etc.—but all principles are based on the only known method to accelerate a particle: to charge it and subject it to an electrical field. This occurs either in one large step or in a number of smaller steps.

Today accelerators are applied in very diverse fields, such as radiotherapy, isotope production, ion implantation, synchrotron light production, spallation, neutron production, radiography, sterilization, inertial fusion, etc., besides of course basic research in nuclear and particle physics. Of more than 15,000 accelerators in operation around the world, only a handful are used in elementary particle physics research, and a few hundred in physics and applied physics research; while one-third are involved in medical applications, such as therapy, imaging, and the production of short-lived isotopes. The vast majority are used for industrial applications, ranging from semiconductor manufacture, electron beam processing, and micromachining, to food sterilization and national security applications, which include X-ray inspection of cargo containers and nuclear stockpile stewardship.

The years around 1930 can be taken as the beginning of the accelerator era, when people at different laboratories started development work employing different principles. Ernest Lawrence and colleagues developed the cyclotron; Robert Van de Graaff, Ray Herb, and others, the electrostatic accelerator; and Cockcroft and Walton, the cascade-accelerator. Gustaf Ising devised and Rolf Widerøe built the linear accelerator. During the 1930s and 1940s the development of accelerators was based on the need from nuclear physicists to obtain higher projectile energies, to exceed the coulomb barrier of heavier and heavier elements for nuclear structure studies, for production of radioactive isotopes, etc. Beginning in the early 1950s, the community of particle physicists also took part in formulating a goal for accelerator development. The first accelerators were built for protons and electrons. Today, ions

R. Hellborg (✉)

Department of Physics, University of Lund, Sölvegatan 14, SE-223 62 Lund, Sweden
e-mail: ragnar.hellborg@nuclear.lu.se

Fig. 1 A modified Livingston diagram showing the exponential growth of accelerator beam energy. The energy of colliders is plotted in terms of the laboratory energy of particles colliding with a proton at rest to reach the same center-of-mass energy
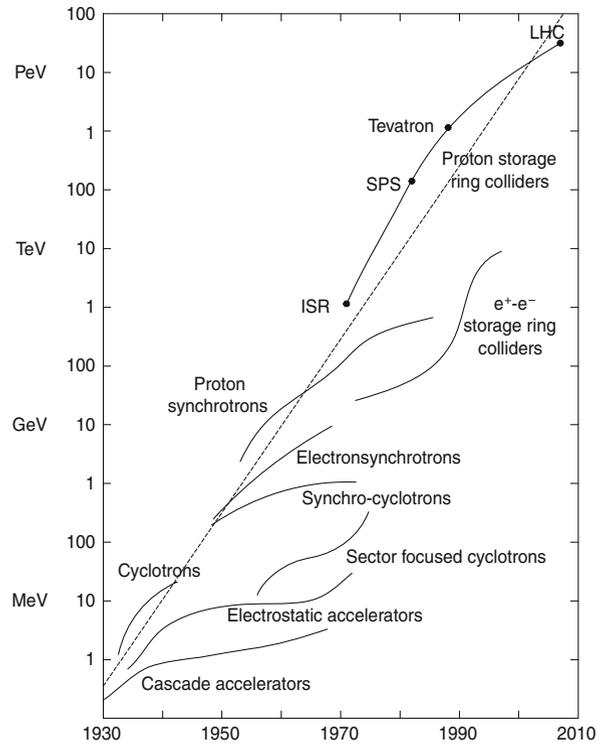
from all elements in the periodic table can be accelerated. Furthermore, it is now also possible to accelerate artificially produced isotopes, i.e., short-lived isotopes far from the stability line, as well as antiparticles like positrons and antiprotons, cluster ions, and low energy, highly charged ions.

The tremendous progress in the construction of accelerators since the 1930s is illustrated in Fig. 1, showing an exponential increase, about an order of magnitude in beam energy per seven years! This graph is called a "Livingston plot" after Stanley Livingston, the accelerator physicist who first constructed such a plot in the 1960s. In the same time period, the cost per eV beam energy has been drastically reduced, roughly by a factor of one thousand. The similarity with Moore's Law in electronics (see chapter "High Speed Electronics") is striking.

2 Accelerators for Low and High Energies

In this section some technical details for accelerators used for nanoscience—cascade accelerators, electrostatic accelerators, and to some extent cyclotrons—are discussed. A more detailed technical description of different types of accelerators can be found in the chapter no. 1 in ref [1].

2.1 Direct Voltage Technique

In accelerators designed using this principle, the particle (after ionization) is accelerated through an accelerating tube in one step. The tube is constructed as a long rectilinear drift tube with a number of electrodes along the axis, and with a controlled voltage for each electrode, partly to aid focusing the beam and partly to distribute the voltage gradient uniformly along the insulation surfaces. The positive ions (or electrons) to be accelerated are generated in a source located at high voltage (except for tandem accelerators—see below—in which negative ions are generated at ground potential). Direct voltage accelerators are often identified with the type of high voltage generator used. The high voltage can be generated by rectifying AC voltage (often called a cascade generator) or by using electrostatic charging, in which a mechanical system carries the charge to the high voltage terminal (called the electrostatic accelerator). Open air accelerators fail above a few MV, mainly because of the moisture and dust in the air, allowing sparks. The Direct Voltage Accelerators are attractive because all types of ions can be accelerated and the ion energy can be changed continuously.

Cascade accelerators. In 1932 Cockroft and Walton built an accelerator along this design and performed the first nuclear reaction study using an accelerated beam of ions. The high voltage unit consists of a multiplying rectifier–condenser system. A principle drawing of one type of high voltage unit of a cascade accelerator is shown in Fig. 2. The main advantage of cascade accelerators is large output currents of up to several hundred mA. This type of generator has for several years been used as a tool in industry for many applications, either as an ion accelerator or as an electron machine; and several thousands of different products are processed with
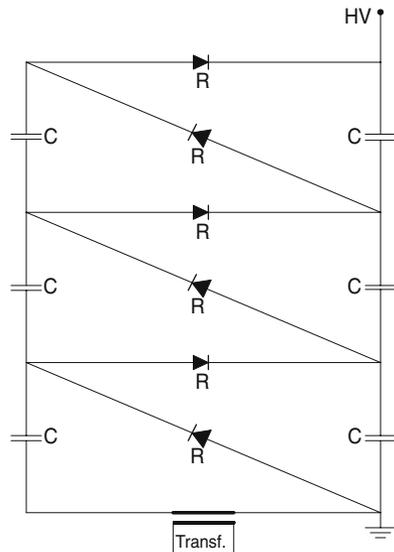
Fig. 2 Asymmetrical circuit cascade generator consisting of capacitors C and rectifiers R

them daily throughout the world. Cascade accelerators are also used as injectors for high-voltage accelerators, because they can provide the necessary high beam current. Cascade accelerators are often used in air at hundreds of kV and below. When enclosed in a pressurized tank, up to 5 MV has been reached.

Electrostatic accelerators. In 1929 Robert Van de Graaff demonstrated the first generator model of this type. An electrostatic charging belt is used to produce the high voltage. Two rollers are provided, one at ground potential driven by a motor and the other located in the high-voltage terminal, well insulated from ground and driving an alternator to power the high voltage terminal electronics. An endless belt of insulating material passes over the rollers. Charge is sprayed from sharp corona points onto the moving belt. The belt conveys the charge to the insulated high voltage terminal, within which the charge is removed by collector points and allowed to flow to the surface of the electrode. In many modern accelerators the insulating belt has been replaced by a chain of metal cylinders. The cylinders are connected to each other by insulating links. By using a metal chain a more robust transport system with much better defined charge transport and hence better voltage stability is obtained. A schematic drawing of an electrostatic accelerator is shown in Fig. 3. Thousands of electrostatic accelerators have been installed around the world. The main disadvantage of the electrostatic accelerator compared to the cascade
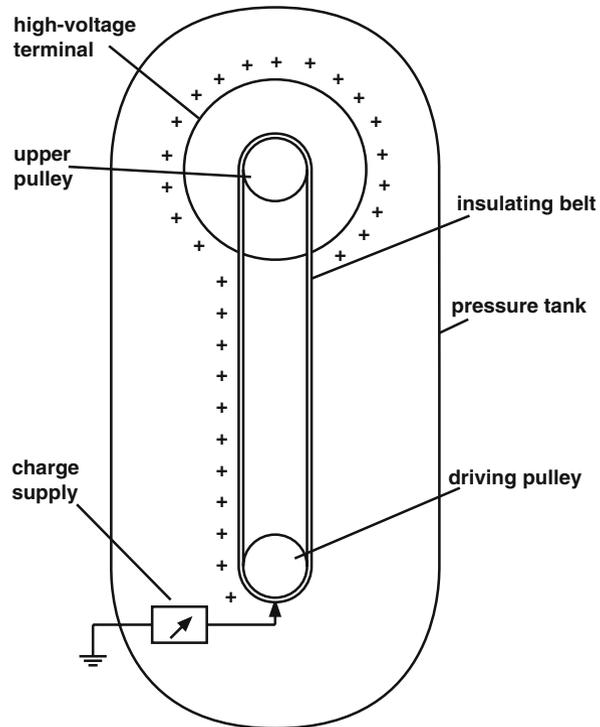
Fig. 3 Principle of an electrostatic generator

accelerator is the low current output. The voltage available today, if the accelerator is enclosed in a tank with a suitable gas of high pressure, is up to a few tens of MV.

Tandem electrostatic accelerators. During the 1950s negative ion sources were developed; beams of ions with an extra electron added (i.e., negative ions) became available. This development made it possible to build two-stage (or tandem) accelerators. The high voltage in these is utilized twice, as can be seen in Fig. 4. Negative ions are formed at ground potential and injected into the first stage, where acceleration to the positive high voltage terminal takes place. The energy gain is eU_T eV, where e is the elementary charge and U_T is the terminal voltage. In a stripper system, a few electrons are stripped off the negative ions in the high voltage terminal, converting them to positive ions. In the second stage, the positive ions once more gain energy as they are accelerated back to ground. Now the energy gain is qeU_T eV, where q is the charge state of the ion. Thus a total energy gain of $(q + 1)eU_T$ eV is obtained. For heavy ions and high voltages, q can be quite high, and therefore the final energy of the ions can be hundreds of MeV.
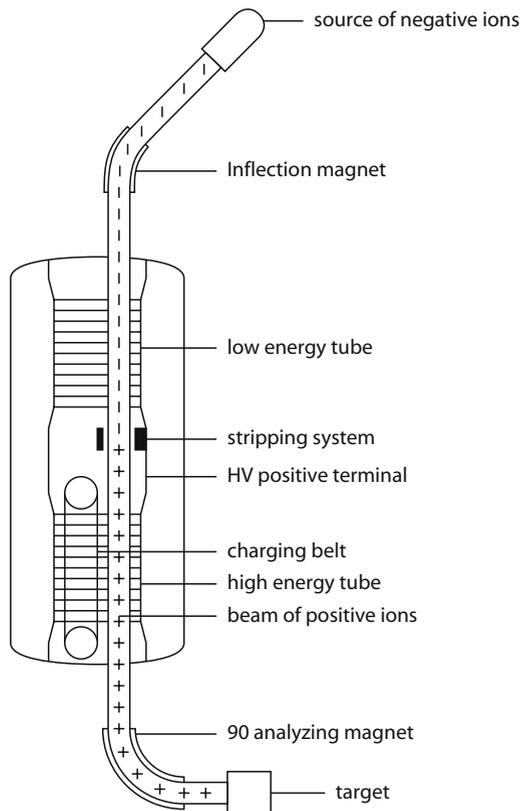
Fig. 4 Principle of a two-stage (tandem) accelerator

2.2 Resonance Acceleration

Another principle is the use of resonance acceleration involving a radio frequency field. The particle has to pass small potential gaps a number of times in resonance with an oscillating electric field. This can be done either in a series of gaps along a straight line, called a linear accelerator, or by a single gap in a circular machine, i.e., cyclotrons, synchrocyclotrons, microtrons, etc.

Cyclotrons. Ernest Lawrence and Stanley Livingston built in 1931 the first demonstration model of a cyclotron for accelerating particles. The principle is illustrated in Fig. 5. Within a flat, cylindrical vacuum chamber placed between the poles of a dipole magnet are two D-shaped electrodes consisting of hollow, flat half-cylinders. The two "Dees" are connected to the RF source (usually working within the range of 10–30 MHz), so that an alternating voltage appears across the gap separating the Dees. An ion source, located at the center of the chamber, produces the ions and supplies them with a low initial velocity. The path of the ions (with the electric charge q and mass m) is circular in the magnetic field B. The radius R of the circle is given by:

$$R = mv/qB \tag{1}$$

Since R is proportional to the velocity v, the period of circulation T (and thereby also the frequency f) is constant for all values of R. Once in resonance, the ion will receive an energy gain each time it passes the acceleration gap between the Dees.

The energies which can be obtained with protons are up to 20–30 MeV, corresponding to a magnet diameter of little more than 1 m. As the beam in a cyclotron travels outwards towards the edge of the magnet, the magnetic field lines are distorted from true straight lines. The curvature of the field lines gives a net force component towards the median plane, which tends to provide focusing counteracting the tendency of the beam to diverge. At the same time the field loses its uniformity and the resonance condition can no longer be maintained.
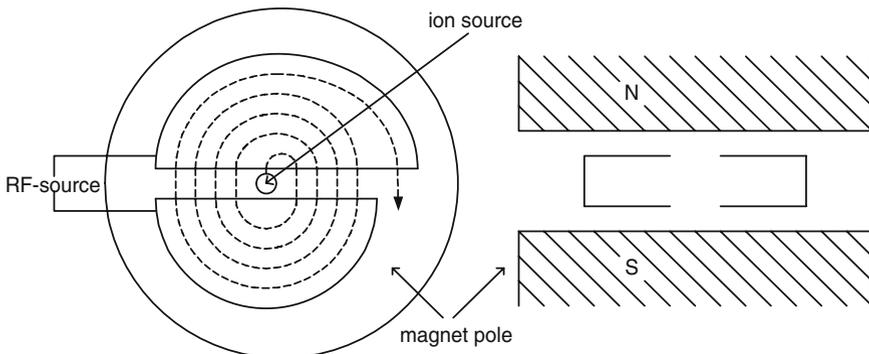
Fig. 5 Principle of a cyclotron

Table 1 Some parameters for accelerators used in nanoscience

	Cascade acc	Electrostatic acc	Tandem el. acc	Cyclotron
Beam current	Several hundred mA	Hundred μA	Tens of μA	Tens of mA
Voltage/Energy range	Up to 5 MV	Up to 10 MV	Up to 25 MV	20–30 MeV
Beam power	Several hundred kW	A few kW	Hundreds of W	Hundreds of kW
Energy stab. $\Delta E/E$	10^{-5}	10^{-4}	10^{-4}	10^{-2}
DC/pulsed	DC	DC	DC	Pulsed
Energy changing	Easily and continuously	Easily and continuously	Easily and continuously	Difficult and not continuously

The advantage of a cyclotron compared with an electrostatic accelerator is that a much higher beam current (tens of mA) is available from a cyclotron. The disadvantages of the cyclotron are that the beam is pulsed, it is difficult to change the beam energy, and normally, this change cannot be made continuously. The energy resolution of the beam is also much worse compared to the electrostatic and cascade accelerators. In Table 1 key parameters are compared for the accelerator types discussed in this section.

3 Ion Optics, Beam Diagnostics, and Computer Control

The ion beam from the ion source of a cascade or electrostatic accelerator is divergent. The source is often followed by an accelerating gap (often the ion source extractor or a downstream gap lens). The next optical unit is the accelerating tube. The tube consists of a (strong) entrance lens, a field region where the acceleration takes place, and a weak exit lens. The accelerator is mostly followed by a quadrupole doublet or triplet. This lens will focus the beam at the object point of a dipole. The magnetic dipole acts as a strong lens to refocus onto the image slits. The dipole is also used to convert variations in beam momentum (magnetic dipole) or energy (electric dipole) into spatial displacements that are detected at slits and fed back to a control system which corrects for errors in the accelerating voltage. The first dipole magnet can be followed by additional lenses. The ability to change beam direction is a valuable utility, and multi-port magnets facilitate directing the beam to more than one target station. Between different dipoles, focusing elements like quadrupoles may be necessary. For low energy (<300 keV), the magnetic analysis of the beam often takes place after the extraction acceleration (at ~ 20–50 kV); analysis is then followed by a second acceleration (or for very low energy, a deceleration) stage. The advantage of this arrangement is that the secondary acceleration supply then only needs to supply enough current for the analysed beam and not the full output of the source, which may contain different charge states, molecular beams, or beams of ions from ion source components (e.g., W from filaments).

The negative beam from the source of a tandem electrostatic accelerator is very divergent and has a low energy. It has to be focused by a lens and accelerated in a first accelerating step. This acceleration can be from a few tens of kV up to several hundred kV depending on the size of the tandem. After this first acceleration, the beam is generally analysed in a magnetic or electrostatic analyzer before being fed to the accelerator. The entrance lens of the accelerating tube behaves like a very strong lens. To avoid the loss of part of the beam, the beam emittance when entering the tube must be within the acceptance of the accelerator. The emittance area of the beam is mainly determined by the design of the source in use. The acceptance area of a tandem accelerator is governed by the dimensions of the stripper canal and the entrance of the low-energy beam tube, in conjunction with the lens effect at the beginning of the tube. For a well designed system, the beam will have a focus in the stripper canal in the centre of the tandem. In the high energy accelerating tube, the beam is divergent and has, after the accelerator, to be handled similarly to the beam from a single stage accelerator.

Focusing of the ion beam along the experimental beam line down to sub-micrometer dimension is described in the chapter "Focusing keV and MeV Ion Beams".

Beam envelope and trajectory calculations can be made today by a number of general purpose beam transport codes. They will, in most cases, describe the real situation in a very satisfactory way.

A number of parameters are related to an ion beam. The main parameters are the type of isotope, charge state, and energy. These parameters are assumed to be determined by the operation of the source and setting of the analyzing magnets and of the different acceleration voltages. The emittance and brightness of the beam are given from the setting of the source and acceleration voltage. In the case of a tandem electrostatic accelerator also, the stripping process will influence these parameters. Very few accelerator laboratories actually measure these parameters continuously. Other parameters—such as beam current, beam diameter, position and intensity distribution, and energy spread—are often measured during operation by different types of diagnostic instruments. Information about the beam current is needed for optimizing the operation and in some types of experiment for calculating the implanted fluence. The current can be measured by a destructive or a nondestructive method. Examples of destructive methods are total beam stopping (with a Faraday cup) and partial beam stopping (with a scanning wire or a rotating sector disc). An example of a nondestructive method is measurement of the residual gas ionization. Monitoring the beam diameter, position, and stability can be achieved by a screen which emits light under beam irradiation. The screen can be a plate coated with a luminescent material like ZnS, MgO, or Al_2O_3. Another possibility is a quartz disk. The lifetime of a coated screen is limited, and they have to be replaced regularly. To suppress discharge effects when using quartz, the irradiated surface can be covered with a metallic net. The emitted light can be observed directly through a glass window or in a remote mode by use of a TV camera. The position and intensity distribution of the beam can be measured by a beam profile monitor. Such a device consists of one or more insulated, thin, moving wires. The current signals from the

wires give information about the intensity distribution in the horizontal and vertical directions. The voltage stability of an electrostatic accelerator (and therefore also indirectly the energy stability of the beam) can be monitored by a capacitive pickup. This is an electrode at the tank wall connected to a current integrator. Since the electrode current is proportional to the derivative of the terminal voltage, the output of the integrator represents the AC component of the terminal voltage. The signal is usually displayed on an oscilloscope and may be fed back into the accelerator stabilization system to improve stability.

Some decades ago, the control of an accelerator and associated elements was carried out through a central console containing a large number of knobs, switches, meters, etc. This console was often located some distance from the accelerator and with long and expensive cables. Each element in the system was controlled by a—often homemade—chassis in the console. The startup of the system, or major changes of the settings, could require help from a skilled operator. Often it required hours of tuning and retuning of the system. In most modern accelerator systems, a properly designed computer system has limited most of these issues. Computerized control systems are less expensive to implement, more reliable and precise, and also easier to modify (see chapter "Electronics for Ion Beam Applications in Nanoscience"). The change from one set of parameters to another is rapid and more flexible. A large number of software and hardware solutions are available for such a control system. Interested readers are referred to, for example, chapter 16 in ref. [1]. A principle design of a control system consists of a computer, a device interface, and the different devices. The operator interacts with software in the computer. The computer communicates with the device interface. This unit contains all the analogue and digital inputs and outputs needed to control the different devices. The devices can be power supplies, vacuum valves, beam stops, profile monitors, etc. In large accelerator systems, several device interfaces can be needed and may require multiple computers distributed throughout the whole accelerator complex.

Several software packages are available. LabVIEW is maybe the most widely used among small and medium accelerators. The communication schemes between the computer and the interface include copper cable, fiber optics, and networks. The choice for controlling devices at elevated potential—like the high voltage terminal or ion source platform—is fiber optic communication, and even for systems located at ground potential, optical isolation is strongly advised.

4 Ion Sources

Below, positive and negative ion sources used for ion beam formation for nanotechnology are described in some detail. Readers interested in more details about positive and negative sources are referred to chapters 11 and 12, respectively, in ref. [1].

4.1 Positive Ion Sources

Electron Bombardment sources: There are many types of ion sources [2], designed to address different needs and for application to different types of accelerators. There are relatively few that have been widely adopted for commercially produced accelerators. These fall into broadly three groups: electron bombardment sources, radio frequency (RF) sources, and microwave sources. The output of these sources ranges from a few μA to several tens or even hundreds of mA. For lower energy accelerators ($<\sim 300$ keV), electron bombardment sources are most common and some designs have been highly developed for applications in the semiconductor industry.

In the context of nanotechnology, common low energy applications are likely to involve ion implantation of metal ions for the production of nanoclusters in insulating or semiconductor materials. Useful transmitted beam currents for a single isotope of metal ions are usually in the 50 μA to 2 mA range, as beam current density on target is typically 1–50 μAcm^{-2}. This range arises because a very low dose rate will often affect nanocluster formation at the lower end, and excessive substrate heating can be problematic at the upper end, depending upon the substrate properties and final energy of the ion, of course. Sources utilize a discharge (40–200 V range, 60–70 V is typical). The beam current is controlled by the discharge current, often (incorrectly) called "the arc," (up to several amperes, 1–2 A is usual) from a directly heated thermionic cathode (filament), or in some variants, an indirectly heated cathode, to the body of the source chamber or anode. The electrons pass through a gas or vapour containing the desired beam element, which is introduced into the discharge chamber at ~0.1 Pa. The electrons emitted by the cathode are accelerated away by the discharge voltage and confined by a permanent or electromagnetic field, depending upon the design. The confinement improves the probability of electron collisions with atoms in the discharge chamber, increasing the plasma density and ions available for extraction. The filament current (and hence temperature) is often used to control the discharge, maintaining the discharge voltage and current at stable values. Examples of this type of source are the Chordis source [2], which emits beams through a circular extraction aperture, and the Freeman source, which was very widely adopted by the semiconductor industry.

The Freeman source (Fig. 6) has a slit aperture, providing a laminar beam aligned with the dispersion plane of the analysis system, to allow the beam to be tightly focused onto the resolving slit and provide very high mass resolution, if required. Both of these sources (and others of a similar type) are available in versions optimized for use with gases (often called cold versions); oven/chemical synthesis variants (hot versions), as shown in Fig. 6; and sputter versions, where the oven is replaced by a sputter assembly.

The gas source is the simplest basic design, producing ions such as inert gases, nitrogen, or ions from gaseous compounds (e.g., CO, AsH_3, PH_3). The oven/ chemical synthesis version is provided with an oven to evaporate elements or compounds into the ion source chamber. To provide a wide range of elements with sufficient vapour pressure for stable operation, a well controlled (sometimes as close as
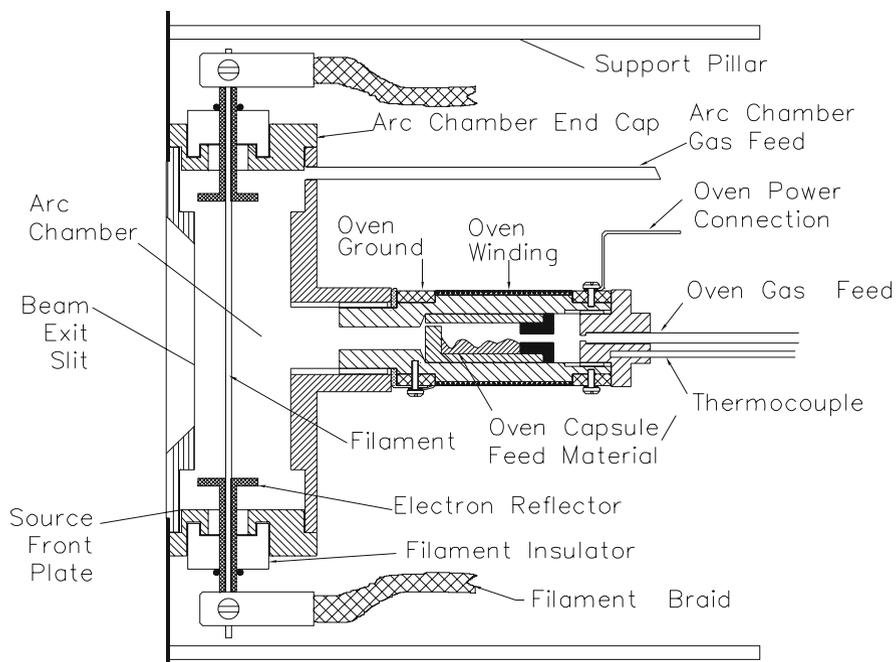
Fig. 6 A sectioned view of the main components of a typical Freeman source in "hot" form. For gas use the oven is omitted, and for sputter use the oven is replaced by the sputter electrode, which projects into the arc chamber through an insulator in the oven port

$\pm 1\,^\circ$C) temperature range to around $1100\,^\circ$C is required, and several oven subassemblies may be needed to address problems such as positive thermal feedback from the discharge chamber or condensation of the evaporated material. Oven sources for research implanters are usually provided with a means to introduce additional gases through the oven; and this enables chemical synthesis of compounds from heated precursors which have vapour pressures that are too low for stable operation at the maximum oven temperature. The reactive gases used are often CCl_4 or Cl, the reaction generating volatile chlorides. This process is sometimes less temperature sensitive than evaporation.

The choice and quality of oven feed material is critical for trouble-free operation of many types of ion source, as is an understanding of the properties of the material. For example, use of a metal which has a significant amount of oxide present may result in rapid evaporation of the oxide at a lower temperature than expected for the metal, temporarily swamping the source and causing premature failure. The solution here is to warm the feed material slowly, releasing the oxide in a controlled manner (often called "conditioning") and/or use material with a minimum of oxide, if possible. Water contamination is often the cause of poor performance and is common with otherwise convenient chloride or fluoride compounds which may be hygroscopic or deliquescent. Tell-tale HCl peaks by the side of the chlorine pair

in the beam spectrum will be present until the water vapour is cleared. Sources using boron nitride components (ovens, filament insulators, etc.) may also produce water vapour from new components when they are heated, unless the material has been previously outgassed in a vacuum oven. In extreme cases the components may even fragment due to buildup of internal vapour pressure. Precautions (e.g., exclusion of air, inert gas storage, etc.) may be necessary to obtain and maintain the dryness of compounds and source components during storage and handling. Condensed hygroscopic or deliquescent compounds in the ion source chamber can also cause vacuum and arcing problems after the vacuum has been broken to change the source, if moist air is not excluded during the process. When reactive gases or compounds are used, a large number of unwanted atomic and molecular beams may originate from the ion source construction materials, or previous feed materials. Sometimes these contaminant beams may not be separable from the desired beam, and in this case it is necessary to identify the contaminant and change the component material, or if possible select a different, less abundant isotope which is better separated from the mass of the unwanted beam.

The sputter source was designed primarily for the platinum group of metals. These elements do not have stable chlorides within the operating range of most sources. However, it is also useful for other medium and high melting point metals (copper, titanium, tantalum, molybdenum, etc.), because the metal is used and is generally simple to handle. The sputter electrode of the desired ion species is introduced into the discharge chamber through an insulator in the rear of the chamber and is biased negatively at up to several hundred volts to attract positive ions from the plasma. The sputter current is used to control the sputtering rate, and the sputter target needs to be cooled to avoid runaway conditions or melting of the electrode at high temperatures. The energetic ions sputter-eject metal atoms from the electrode into the plasma on impact with the sputtering target. The primary discharge is often argon, but using higher mass noble gases can increase the sputter yield (and cost). Sputter versions of electron bombardment sources usually produce substantially less beam current than the gas or oven variants, often a few hundred μA, because a relatively small percentage of the sputtered atoms are ionized, with the remainder inevitably condensing on the walls of the discharge chamber. Flaking of the condensed metal is a common cause of short circuits within these sources, usually after the shutdown/startup thermal cycle. For nonplatinum group elements with relatively high melting points, the introduction of CCl_4 or Cl to the argon sputtering discharge can significantly increase the beam current (sometimes by as much as a factor of 10 or more) and reduce condensation by creation of more volatile chlorides, which can evaporate from the hot inner walls of the discharge chamber. Precise reactive gas control is necessary for stable operation, particularly for elements which have vigorous reactions with chlorine.

Electron bombardment sources produce predominantly singly charged ions. They tend to be designed to produce stable discharges with a minimum of noise (often termed "hash") and relatively few multiply charged ions are produced. Typically 5% of the singly charged beam may be doubly charged and 0.5% triply charged. These proportions may be increased to some extent by careful tuning and running

the source aggressively with very high discharge voltages (150–200 V), but stability suffers, with hashy beams and rapid filament erosion due to high arc voltage, and the source lifetime may be severely affected.

In general, the lifetime for electron bombardment sources between services is in the 20–50 h range, which is usually not a problem in a research lab, but in production environments there is pressure to extend this to 100 h or more.

Ions are extracted from the anode of the source by biasing it positively with respect to a ground electrode. This requires all the ion source electronics to be supplied with power via an isolation transformer. A negatively biased electrode (typically 1–5 kV, sometimes termed suppression electrode), is often placed between the source outlet and ground electrode to prevent secondary electrons from beam interactions backstreaming to the source. The geometry of the outlet plate and extraction system is critical for good performance, and in some cases is provided with a variable gap between the source and suppression/ground electrode pair, so that the strength of the electric field can be increased at low extraction voltages. The transmitted beam current is proportional to $U^{3/2}$ (where U is the applied voltage), from the Child-Langmuir Law [3, 4]. For this reason, some systems are designed to extract and transport the beam at higher energy, then decelerate the ions close to the target for very low energy beams (< 10 keV). However, this approach can also bring problems, as any energetic neutrals will not be decelerated.

RF Ion Sources. Radio frequency ion sources are used on a wide range of accelerators from low energy (20–200 keV) cascade machines, such as neutron generators, through to much higher energy electrostatic accelerators for the production of ions from gaseous feeds (H, D, He, N, etc.). The gases are usually stored in small lecture-sized cylinders on the high voltage terminal of the machine and are "leaked" into the source bottle near the base via a controlling device called a thermomechanical leak, although there are other methods (e.g., palladium leaks for hydrogen and motor-driven needle valves). High purity gases are used and care is needed to purge the pipework, as some contaminants (nitrogen in particular) result in disproportionately high beam fractions, due to lower ionisation potentials.

The body of the source is constructed from a glass tube which is closed at one end with a metal extraction pin fused into that end. In front of the extraction electrode is a silica baffle plate with holes around the perimeter. This baffle is there to prevent the extraction voltage arcing through the plasma. The other end of the tube is bonded to the metal source mounting plate, which also mounts the metal extraction canal (which is typically 1–2 mm diameter), and the connection for bleeding the gas into the bottle. The canal projects into the source and is insulated by a silica boron nitride or alumina sheath. The RF, which can range from 10 MHz up to a few hundred MHz, is applied either via a coiled antenna around the bottle or by a pair of coupling rings clipped onto the bottle, and is tuned by a shunting capacitor. The plasma density is increased in the region of the canal by means of a solenoid coil around the base of the source bottle. The pressure in the source bottle is usually on the order of 0.1 Pa, but the pressure in the accelerator tube is necessarily much lower, which is achieved by differential pumping through the small canal. The beam current is controlled by variation of the extraction voltage in combination with gas pressure adjustment.
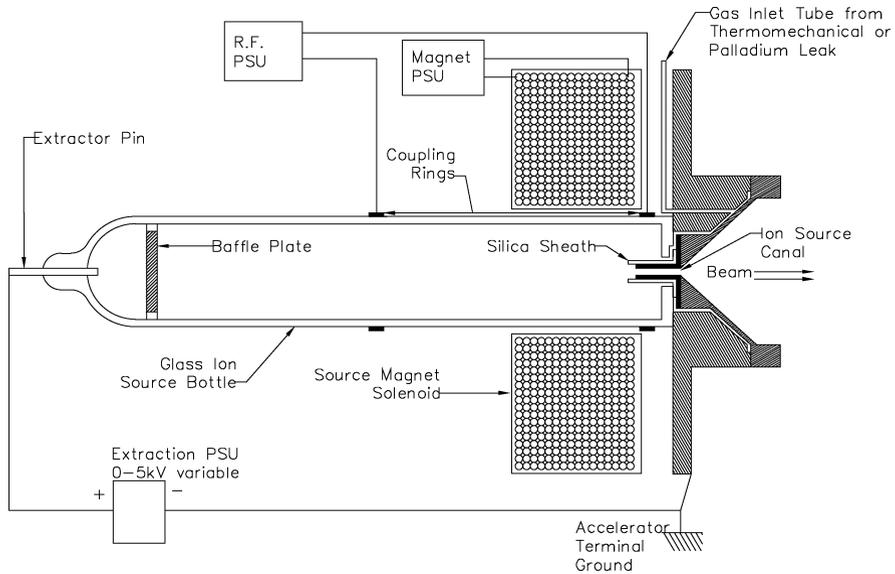
Fig. 7 A typical RF source for gaseous ions used on many single-ended accelerators

This type of source will normally produce up to a few mA of beam and works well for light ions where electron bombardment sources do not perform efficiently. In terms of nanoscience they are most likely to be employed on analytical accelerators or for coimplantation of light ions. A basic RF source similar to those used on many electrostatic accelerators is shown in Fig. 7 [5].

ECR ion sources. The electron cyclotron resonance (ECR) source is a type of microwave source, which has applications mainly as a cyclotron ion source as well as for materials processing. The plasma from a gas feed is excited by a microwave source typically in the 5–15 GHz range, and the electrons are constrained by an axial and multipole magnetic field, arranged to make them pass many times through the plasma. Consequently, these sources can produce an excellent yield of highly charged ions, because of the high electron temperature. The absence of a consumable cathode enables the source to run with excellent stability at high power for long periods of time, and the selective heating of the electrons by electron cyclotron resonance heating means that the energy spread of the ions is very small. Beams of ions from solid materials can be produced by direct insertion of the material into the plasma for high melting point materials, or from an oven for lower melting point species.

4.2 Negative Ion Sources

Most elements and many molecules can form negative ions by adding an extra electron, with lifetimes long enough to be used in accelerators. The negative ions can be formed through a number of processes. These processes are all exothermic. The

binding energy of the additional attached electron is the parameter that characterizes the negative ion state. It is called the electron affinity and must have a positive value to form a bond state. About 80% of the elements have a positive electron affinity. Unfortunately, a few have negative electron affinity, e.g., Be, Mg, N, Mn, the noble gas atoms, and a few others.

One type of negative ion formation is in a plasma environment. Those sources employ electron impact attachment methods to form negative ions directly in a plasma and are often called direct extraction sources. Sources belonging to this group are the duoplasmatron-type source, diode negative ion source, and Penning discharge source. The materials to be ionized are fed into the plasma as gases, vaporized from an oven, or sublimed from a solid rod of the material.

A second type of negative ion formation is based on charge exchange. During a collision between an ion or a neutral atom/molecule and another ion or neutral atom/molecule, an electron can be transferred from one particle to the other. The probability depends on collision energy, ionization potentials, and electron affinities of the colliding partners. If a positive beam is to be converted to a negative, each ion has to undergo two collisions, first to become neutral, and then to become negative. The practical arrangement is often a positive ion source followed by a charge-exchange cell. The source can be a gas source, such as a duoplasmatron, or an rf-source, or if needed, a more universal source capable of producing positive ions from solid material. Several materials, including Li, Na, Mg, Ca, Rb, and Cs, have been utilized as charge-exchange media. Efficiencies over 90% have been obtained for suitable combinations between beam and charge-exchange media.

A third type of negative ion formation is through thermodynamic equilibrium surface ionization. Atoms (and molecules) impinging on a hot metal surface can be emitted as positive or negative ions. Unfortunately, negative surface ionization has not been utilized frequently for generation of an ion beam, owing to the lack of chemically stable low-work-function material for use as ionizers. However, for positive surface ionization, this technique is used frequently, because several high-work-function metals may be chosen for that purpose. LaB_6 is normally used for negative surface ionization.

A fourth type of negative ion formation is through nonthermodynamic equilibrium surface ionization. This is the most widely used technique today to obtain a negative beam. A surface, covered with a thin layer of a highly electropositive absorber like Cs, is sputtered by a primary beam. The sputtering process is measured in terms of the yield of ejected surface atoms per incident primary particle. This is a nonthermodynamic equilibrium process. The positive ions in the primary beam are formed either by direct surface ionization of a group IA element (Cs, Rb, etc.) or in a Cs-rich noble gas plasma discharge. The primary beam is accelerated to between a few hundred eV and several keV, to sputter a sample containing the element of interest. A small proportion of the sputtered atoms is negatively charged and is thus accelerated from the negatively biased sample through an aperture.

A few typical negative ion sources of different design principles are described below in some more detail.

Duoplasmatron. This type of source belongs to the group "negative-ion formation in a plasma environment." Duoplasmatron sources for production of positive ions were in use long before their capability for negative production was discovered. It was found that negative ions are more abundant in the periphery of the plasma, and by offsetting the extraction electrode with respect to the centre of the plasma, the yield can be enhanced. To extract the negative ions from the source, it needs to be operated in reverse polarity. A duoplasmatron can produce $>10\mu A$ of H^- and D^- from H_2 and D_2 gases, respectively. A few heavy ions (e.g., C^-, CN^-, O^-, F^-, S^-, Cl^-) can be produced from gaseous or liquid species and with currents >1 μA. When producing heavy ions, the duoplasmatron often becomes contaminated in a short time and the lifetime of the filament is decreased dramatically. This happens even if the gas feeding the source contains only a few percent of the gas/liquid (like CO_2) added to the H_2 to obtain the heavy ions (like C^- or O^-). A negative duoplasmatron source producing a hydrogen beam has an emittance area of $(1-2)$ π mm $\cdot$ mrad $\cdot$ $MeV^{1/2}$ for nearly 100% of the beam.

Charge-exchange cell. The combination of a positive ion source and a charge-exchange cell for converting the positive ions to negative can, as mentioned above, have a very high efficiency. For suitable combinations of ion type–donor material–ion speed, it can be over 90%, i.e. nearly all positive ions will be converted to negative. For heavy ions in particular, a very high efficiency can be obtained. To obtain negative helium ions, the charge-exchange method is the only one available (as He^- is a metastable state). Unfortunately, the efficiency for He^- production (for all donor material) is very low, around one percent. To maximize the negative current and to increase the operational lifetime, special attention has been given to the cell design. By accurately modeling the thermal gradients in relation to the physical design and the construction material, the lifetime of the cell can today be several weeks to many months. The cell should be designed with temperature gradients along the entrance and exit pipes to ensure that the donor vapour condenses, liquefies, and drains back into the reservoir. The main problem with earlier cells was the condensation of the donor material in the entrance and exit openings. The negative beam currents from a cell are strongly dependent on the type of positive beam and donor material, as well as on the beam energy. The emittance of the negative ion beam is in the range of $(1.6-2.8)$ π mm $\cdot$ mrad $\cdot$ $MeV^{1/2}$ [6]. In Fig. 8 a schematic drawing of a positive duoplasmatron and a charge-exchange cell is shown.

Spherical geometry ionizer sputtering source. This source belongs to the group "nonthermodynamic equilibrium surface ionization." The main development of sputtering sources started around 1970 and has gone through several development stages. Most tandem accelerators around the world are today equipped with sputtering sources, and many of them are supplied with an ionizer (to obtain the Cs^+ ions) with spherical geometry. Both computer simulations and experimental experience show that the sputter sample is optimally positioned at the focal point of the ionizer. For applications such as accelerator mass spectroscopy (AMS), where high efficiency or high-frequency sample changes are desirable, the source has been designed with a multisample wheel. Typical beam currents are tens of μA, with exceptions for some ions having higher/lower currents. An emittance area of

Fig. 8 Schematic representation of a close-coupled positive duoplasmatron and a charge-exchange cell, from [1]
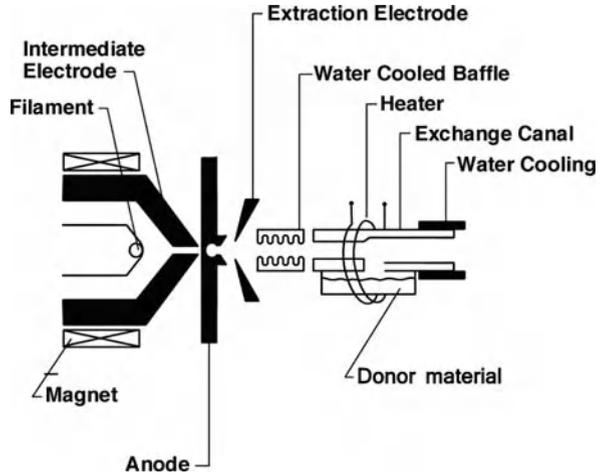
$(5–7)\pi$ mm $\cdot$ mrad $\cdot$ MeV$^{1/2}$ is typical for this type of source. In Fig. 9 a schematic drawing of a spherical geometry sputtering source is shown.
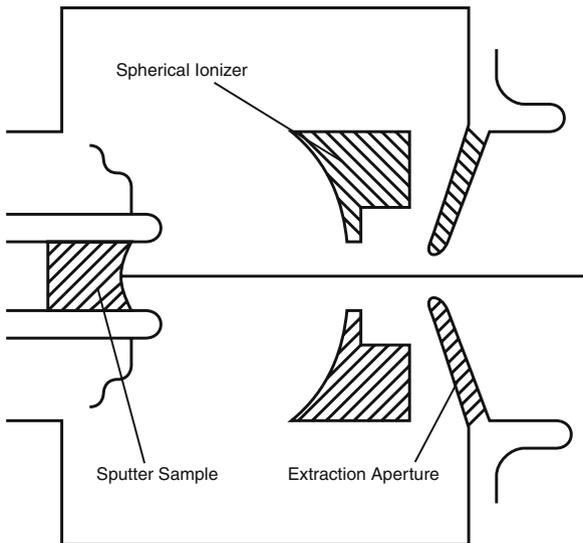
Fig. 9 A Cs-sputtering ion source equipped with a spherical geometry surface ionizer, from [1]

5 Requirements for Use in Nanoscience

The low energy applications (< 250 keV) of accelerators are dominated by ion implanters and most of the remainder of this chapter relates mainly to the post

beam–analysis hardware which is more specific to implanters. Before the beam analysis, implanters are similar in principle to other accelerators.

As applications drive towards smaller and smaller scale devices (e.g., as discussed in the chapter "High Speed Electronics"), lower energies are often needed. Production and transport of low energy beams is difficult, and so either beams are extracted and transported at higher energies and then decelerated close to the target, or molecular beams can be used; or for very low energies, cluster-ion beams may also be applied. Cluster beams give quite different conditions to monomer implants, because of the overlapping ion tracks of the cluster atoms and sputtering differences. Low energy, highly charged ions can also be useful in the study of nanoscale surface interactions.

5.1 Target Systems

The wide-ranging accelerator systems which have been developed for experimental applications are often constructed or adapted by the users to accommodate the requirements of a particular application. There are relatively few manufacturers of machines, and they are directed mainly at production or ion beam analysis markets. Often it will be necessary to adapt the equipment, particularly the target system (end station) for the needs of new experimental work.

For ion implanted nanocluster applications, there may be several parameters which require quite precise control. This type of work normally implies the implantation of insulators or semiconductor materials; and the relatively high doses required (nanoclusters are not usually seen at fluences below 5×10^{15} ions cm^{-2} [7, 8], as there are insufficient implanted atoms to cluster) and the current density related effects, mean that relatively high beam power is often applied. In most machines the beam profile will not be completely uniform and will change significantly in response to a number of interactive parameters, such as ion source conditions, extraction/acceleration potentials, control vanes, magnet settings, etc. This means that to provide uniform implantation, the beam or the sample must be rastered. For low beam currents, electrostatic beam scanning may be used, but at high currents, magnetic beam scanning or mechanical scanning of the target is necessary. The electrostatic method strips the high current beam of the low energy electrons which maintain space-charge neutralization, causing the beam profile to expand, a particular problem for lower energy, high mass ions. Magnetic and mechanical scanning tend to be lower frequency than electrostatic, and this, combined with the way that the beam is moved over the sample, can give rise to thermal nonuniformity on the sample, even when the dose uniformity is well controlled. For some ions (silver, copper, and gallium in particular), temperature variations as small as 5 °C over the surface of the sample can measurably affect the formation of implanted nanoclusters and hence the optical properties of the sample. This sensitivity means that sample holder design, cooling/heating, and temperature control are critical, especially for insulators. Samples should be bonded (one of the commercially available silver

pastes works well, provided ample time is allowed for drying between impervious materials), or clamped using a vacuum compatible heat transfer medium behind the sample, such as exfoliated graphite paper. Samples which are large, thick, or have poor thermal conductivity obviously present more of a challenge.

5.2 Charge Control and Vacuum

Insulators present challenges related to charge buildup on the surface of the sample. If steps are not taken to neutralize the positive charge resulting from the beam and secondary electron emission (some materials have secondary electron coefficients as high as 10), then a voltage of several tens of kilovolts can build up on the sample, if the beam energy permits. The effect of this voltage is to distort the beam shape at the target, causing spark damage to the sample, and to change the effective energy of the beam due to deceleration. There are several ways of dealing with this problem. The simplest is to ensure that the beam is always larger than the insulating area of the sample and so is never isolated on it; or, for large samples, to raise the pressure (~ 6–9×10^{-4} Pa) by means of a controlled argon bleed and differential pumping near the sample, thus providing secondary electrons from beam interactions with the Ar. Care is needed, because if the pressure is too high ($> 1 \times 10^{-3}$ Pa), there will be significant dosimetry errors due to beam neutralisation. The process may be checked by simply varying the pressure and monitoring the effect on the measured beam current. On the subject of vacuum, care is necessary to maintain the pressure at optimum levels throughout the accelerator, especially for low energy, high current beams. This is because too low a pressure will bring space-charge problems causing blowup of the beam, and too high a pressure causes neutralization of ions and dissociation of molecular beams. Other methods for charge control are placing a filament close to the target to provide thermionic electrons, electron flood systems, or plasma flood systems. All require careful design and controls to avoid dosimetry errors. Care is necessary to avoid energetic electrons reaching the target, as these damage sensitive targets or can generate secondary electrons which exacerbate the problem.

Accurate dosimetry is important and is often provided by measurement of the beam on a beamstop between scans. This method is generally slightly less accurate than the common alternative of making the entire target system a Faraday cage, but is much simpler to implement. Beamstop design is important to ensure full suppression of secondary electrons, which if allowed to escape, result in dosimetry inaccuracy; and also to avoid target contamination from sputtered beamstop materials. Graphite is often the preferred material for low energy applications, and the assembly is usually cooled. Where cooling is used, it is necessary to use either deionized water with a suitable length of insulating tube, or other coolants to avoid dosimetry errors due to current leakage.
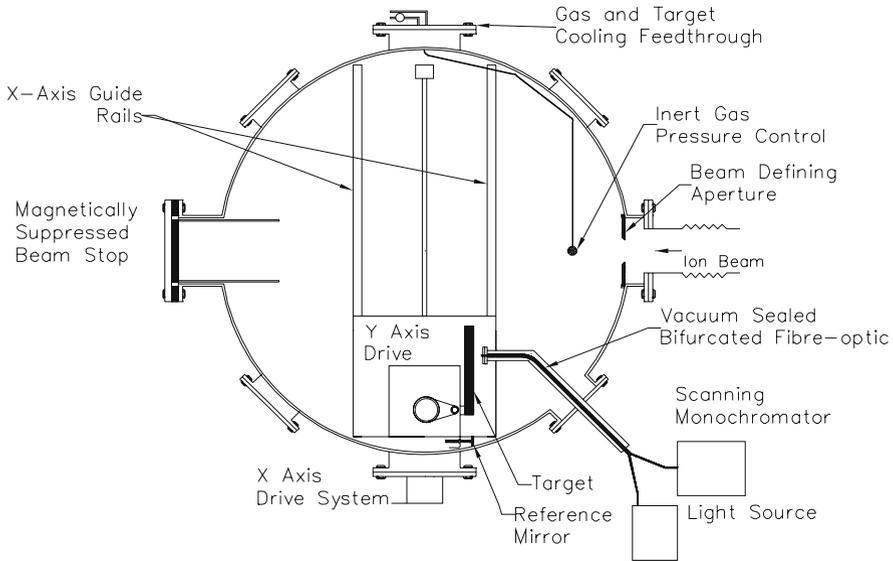
Fig. 10 A plan view of a target system for in situ measurement of sample reflectance. The X–Y scan positioning allows alternate implantation, reference measurement, and sample measurement. The diameter of the chamber is 900 mm, allowing a sample size of 200 mm in the X-plane. The Y size is dependent upon the height of the chamber

5.3 In Situ Measurement

As mentioned earlier, nanocluster formation can be very sensitive to implantation conditions, and in situ measurement systems can be an invaluable aid to analysis, problem solving, and process control. Mass spectrometer vacuum gauges are valuable for documenting and maintaining the quality of the system. For implantation of insulators, in situ measurement of reflectance [7, 8] between dose increments can provide a valuable insight into dose, temperature, and energy dependence experiments. This is because small (or sometimes large) changes, which might be disregarded as experimental error in individual experiments, can be detected; and a single experiment can provide the same data as many discrete implants. An experimental setup for in situ measurement of reflectance is shown in Fig. 10. In principle, a similar technique for transmission measurement is possible, but target holder design which ensures thermal uniformity and control is very difficult. In situ measurement of sample resistivity (between scans) or ion beam–induced luminescence can also be useful in some instances. It should be noted that some sample properties can drift on a scale of minutes or hours after implantation ceases. Also, the interruption of an implant due to equipment problems or time constraints may significantly affect the outcome because of thermal or dose rate effects.

References

1. R. Hellborg (editor). *Electrostatic Accelerators – Fundamentals and Applications* (Springer Heidelberg, 2005)
2. I.G. Brown (editor). *The Physics and Technology of Ion Sources* (John Wiley & Sons, 1989)
3. C.D. Child. *Phys. Rev. (Ser1)*, **32**, 492 (1911)
4. I. Langmuir and K.T. Compton. *Rev. Mod. Phys.*, **3**, 251(1931)
5. C.D. Moak, H. Reese, and W.M. Good. Nucleonics, **9**(3), 18 (1951)
6. P. Tykesson. *Symposium of Northeastern Accelerator Personnel* (Oak Ridge, Tennessee, 23–25 October 1978).
7. L.C. Nistor, J. van Landuyt, J.D. Barton, D.E. Hole, N.D. Skelland, and P.D. Townsend. Colloid size distributions in ion implanted glass. *J.Non-Cryst. Solids*, **162**(3), 217–224 (2 September 1993)
8. D.E. Hole, P.D. Townsend, J.D. Barton, L.C. Nistor, and J. Van Landuyt. Gallium colloid formation during ion implantation of glass. *J. Non-Cryst. Solids*, **180**(2), 266–274 (January 1995).

Focusing keV and MeV Ion Beams

Mark Breese

1 Focusing Different Kinds of Charged Particles

It is worth briefly considering why different focusing systems are used for electron microscopes which typically focus 20 to 200 keV electrons, focused ion beam (FIB) systems which typically focus 3 to 30 keV ions from Be^+ to Ga^+ to Au^+, and nuclear microprobes which focus MeV light ions such as protons or He^+.

Consider the Lorentz force $F_B = Bqv$ on a particle of charge q, mass m, and the nonrelativistic energy T, in a magnetic field B. The resultant acceleration towards the centre of a circular path of radius ρ gives:

$$Bqv = \frac{mv^2}{\rho} \tag{1}$$

Rearranging gives $B\rho = mv/q$. The magnetic rigidity of the beam, $B\rho$, equal to the momentum per unit charge, is a measure of how difficult it is to bend charged particles through a given radius in a magnetic field. This relationship may be rewritten as: $B\rho = \sqrt{2mT}/q$, where the dependence of the magnetic rigidity on both the particle mass and energy is more explicit. The ratios of magnetic rigidities of three singly-charged particles commonly used in the different types of microscopes, 20 keV electrons, 20 keV Ga^+ ions, and 2 MeV protons, are 1:356:426.

A much weaker magnetic field is obviously required to focus electrons, compared with similar energy ions. Axially-symmetric magnetic solenoid lenses have a weak focusing action, since they rely on the lens fringe fields. Since a weak magnetic field is sufficient, they can be used with great effect to focus electron beams down to spot sizes of nanometers in a scanning electron microscope (SEM) with a typical working distance of 10 mm, and to provide resolutions of about 0.1 nm in a

M. Breese (✉)

Physics Department, National University of Singapore, Lower Kent Ridge Road, Singapore 117542, Singapore

e-mail: phymbhb@nus.edu.sg

transmission electron microscope (TEM) [1, 2]. These lenses can be manufactured with near-perfect field symmetry and have intrinsically low chromatic and spherical aberrations. Residual spherical aberrations may be compensated for using octupole correctors [3], enabling resolutions of less than 0.1 nm. Comprehensive descriptions of electron and ion optics are given in [4–8].

2 Focusing keV Ion Beams

FIBs are used in many applications, such as nanoscale milling and sputtering of microstructures, metal layer deposition using a gas injection system, three-dimensional nanofabrication, device modification, TEM sample preparation, elemental microanalysis, etc. [8–10]. Modern FIBS may have a built-in scanning electron microscope, providing the capability of monitoring applications with real-time imaging. Due to the wide scope of applications, the specifications required of a modern FIB cover a wide range. They typically require a beam energy of 1–30 keV containing a focused beam current of 1 pA–10 nA. The beam current density may be as high as $10\,A/cm^2$, and the working distance of the microscope is typically 20–70 mm.

Consider the electric force $F_E = qE$, on charged particles in an electric field E, which produces an acceleration towards the centre of a circular path of radius ρ, i.e.,

$$qE = \frac{mv^2}{\rho} \tag{2}$$

Rearranging gives $E\rho = 2T/q$, so the electric rigidity, $E\rho$, is equal to the energy per unit charge carried by the beam and is a measure of how difficult it is to bend charged particles through a given radius in an electric field. Since electric rigidity depends only on particle energy and not on mass, moderate electric field strengths used in cylindrically-symmetric electrostatic lenses can be used to focus keV heavy ions across the periodic table, whereas conventional magnetic solenoid lenses are not strong enough.

FIBs thus typically use electrostatic condenser and objective lenses for beam transport and focusing, such as cylinder lenses, einzel lenses, or aperture lenses. These comprise two or three electrodes of alternating polarity which decelerate/accelerate the beam as it passes through, resulting in a net focusing effect and no net change in beam energy. Such lenses have low aberrations, and when used to focus ion beams extracted from a high brightness liquid metal ion source, very high resolutions of a few nanometers are attained. Some FIB operating modes require the best possible beam resolutions, whereas others require the highest possible current density. These options require different beam trajectories through the lens system, as shown in Fig. 1.

Fig. 1 Ion trajectories in a FIB for (**a**) normal operation for optimal spatial resolution, and (**b**) producing the highest current density on the sample surface with a beam cross-over between the two lenses
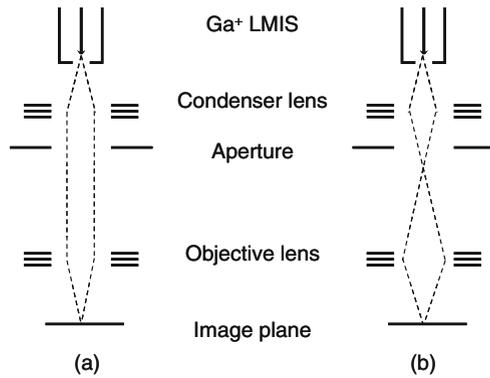
3 Focusing MeV Light Ion Beams

Cylindrically-symmetric electrostatic lenses cannot be used for focusing MeV ions because of their high energy, and magnetic solenoid lenses cannot be used because of the high ion momentum. Magnetic quadrupole lenses, which have a much stronger focusing action because their magnetic field lines are perpendicular to the ion trajectories, are thus used in the majority of microprobes. Drawbacks of using quadrupole lenses include greater lens complexity and difficulties of construction to the required precision. Since each quadrupole lens focuses in one plane and defocuses in the orthogonal plane, more than one lens is needed to form a point focus. Most microprobes use combinations of 2, 3, or 4 quadrupole lenses with alternating polarities and different strengths to focus MeV ions to a small spot. General accounts of quadrupole optics can be found in [11–16], and those specific to nuclear microprobe probe-forming systems in [17–19]. Only more recent work on focusing high-energy ion beams to small spot sizes is discussed here.

Advances in focusing MeV ion beams have mainly resulted from the design and construction of magnetic quadrupole lenses largely free of parasitic harmonic fields which can limit resolution. Such lenses are constructed from a single piece of high-quality soft iron by a wire-cutting "spark erosion" process which is capable of micron precision. The original lenses fabricated in this manner [20, 21] were 10 cm long with a bore diameter of 15 mm. Demagnifications up to 80 were attained, which resulted in beam spot sizes down to a few hundred nanometers. These lenses are still widely used in many microprobes today. Higher demagnifications were not achieved with these lenses because they were designed to focus a beam to an image distance of about 15 cm downstream of the final lens. The long lens length, large bore diameter, and excitation coils which protrude several centimeters beyond the physical lens length limited their ability to focus to a shorter image distance with higher demagnification.

A smaller lens was designed using the same process to overcome these limitations. This new lens achieved a similar focusing strength with a lens body of 5 cm

in length and a 7.5 mm bore diameter [22]. The new lenses were asymmetric, with a 'snout' on one side so that the pole pieces projected 15 mm beyond the main lens body. The purpose of this projecting snout is to move the effective exit face of the final lens as close as possible to the wall of the target chamber in order to reduce the image distance. Detailed field calculations were performed to ensure this would not introduce field saturation within the lens; they demonstrated that the field strength in the bore of the projecting snout was almost the same as within the main lens body. These lenses achieved demagnifications up to 200 with image distances of 5–7 cm, resulting in beam spot sizes of 30–50 nm [23, 24].

Magnetic quadrupole lenses with slightly different designs have also been produced [25]. An orthomorphic "Russian" magnetic quadruplet system is used by another group to consistently achieve resolutions of about 100 nm [26]. Other groups use different approaches to focusing MeV ions. In order to focus heavier or higher-energy ions, a microprobe has been constructed with elongated magnetic quadrupole lenses to increase focusing strength, though at the expense of a lower demagnification [27]. Other groups have tried focusing high energy ions with electrostatic quadrupole lenses which give a strong focusing effect. There have also been attempts at constructing achromatic quadrupole lenses [28, 29], comprising both electrostatic and magnetic quadrupole fields. In principle, these result in the beam focus being insensitive to energy changes and therefore free of chromatic aberrations. However, the electrostatic quadrupole lens requires a very small bore diameter in order to attain sufficiently large field strengths. This makes it difficult to accurately align the individual pole pieces with each other, which tends to introduce large parasitic harmonic fields [30–32]. Other groups have developed more complex quadrupole focusing systems, such as a spaced quadruplet [33], a five-lens system [25, 34], a six-lens system comprising two separated triplets [35], etc. While these systems may have larger demagnifications, it remains to be seen whether any can produce a beam resolution which surpasses, or even matches, that produced by the simpler triplets and quadruplets.

Other approaches to focusing high-energy ion beams include a superconducting solenoid lens [36, 37]. This makes use of the lower intrinsic aberrations of an axially symmetric lens compared with a quadrupole lens, though with increased complexity of design and operation. A different approach to the use of high-energy heavy ions as a lithography tool is to form a projection lens system [38].

3.1 Comparison of Different Lens Configurations

A doublet (two-lens system) is the simplest system, but produces highly asymmetric demagnifications. A quadruplet produces an orthomorphic focus (equal demagnifications in orthogonal planes), but at the expense of the lowest demagnification. A triplet has one positive demagnification because of the beam cross-over within the lenses, and a smaller asymmetry in the demagnifications than a doublet. It gives the highest mean demagnification, so is capable of achieving the smallest beam spot size for a given beam brightness and current. This system is called the 'high-excitation'

triplet because there are other possibilities of forming a point focus with lower lens currents.

In keeping with its higher demagnifications, the chromatic and spherical aberration coefficients for the triplet are higher than for the other two systems. The higher demagnifications of the triplet produce smaller beam spots only when the effect of its large aberrations do not limit the beam spot size. For the triplet, the beam spot size increases rapidly with larger beam divergence angles into the lenses, because of the large spherical aberrations. The quadruplet is the most 'forgiving' of systems, having the lowest aberrations. Similar conclusions are reached for excitation, rotation, and translation aberrations of the quadrupoles; the triplet is generally the most sensitive, so requires the most careful setting up in order to take advantage of its higher demagnification.

3.2 Accelerators

Another vital aspect of attaining highly-focused beams using a nuclear microprobe is to first produce a beam which is as monoenergetic as possible. Many microprobes are based on electrostatic accelerators which are mainly used for conventional ion beam analysis, in which an energy spread of ∼1 keV or more is acceptable. One consequence of such beam energy fluctuations is that the beam position continually moves when it is transmitted round an analysing magnet, so the beam current passing through the microprobe object aperture and consequently into the chamber fluctuates. Such fluctuations degrade microprobe performance in two ways. First, they limit the attainable beam spot size because of the chromatic aberrations of the quadrupole lenses. Second, the fluctuating beam current makes it difficult to uniformly irradiate a given area, which introduces noise into the recorded images for nuclear microscopy or nonuniform irradiations for proton beam writing. This can be alleviated by averaging the irradiation over a period of time, but such intensity fluctuations are still a major limitation to uniformly irradiating polymers and semiconductors with a specific dose. A new generation of megavolt electrostatic accelerators, the Singletron and Tandetron [39, 40], achieve very stable terminal voltages of typically tens of volts, i.e., at least ten times more stable than was previously attainable. The beam current hitting the sample in a microprobe mounted on a Singletron accelerator is thus very stable, with fluctuations as low as 2% reported when the accelerator is properly optimized [41].

4 Other Limitations to Microprobe Resolution

Many limitations to microprobe focusing performance other than intrinsic and parasitic aberrations of the lenses and accelerator voltage fluctuations have been identified over the years and have been solved on many systems. These include:

Thermal drifts and expansion. Magnetic quadrupole lenses gradually heat up throughout the day, which can be a source of long-term drift requiring the beam to

be periodically refocused. High demagnification systems tend to suffer more from this because the lenses tend to be thinner and require high excitations to produce short image distances. A temperature-controlled environment and a fan directing air over the lenses help to alleviate this problem. It can be reduced by using low beam energy and light ions to minimize the quadrupole lens currents needed for focusing. Another source of thermal drift is any lamp used to illuminate the chamber interior. Even a few watts of power can provide enough radiated heat to cause a sample stage to expand by micrometers. This can be reduced by illuminating the chamber with a lamp mounted outside the chamber, or with a fiber optic bundle which provides little heating effect, or with high-efficiency low power light emitting diodes.

Stray magnetic fields. This problem should be taken into account in the construction of a new microprobe. Any source of electromagnetic noise close to the beam line can degrade microprobe performance by continuously shifting the position of the focused beam in the chamber. All sources of noise, such as monitors, transformers, and vacuum gauges, should be moved as far as possible from the beam line to minimize this problem. The simplest means of identifying any noise source is with a search coil connected to an oscilloscope. Reference [42] describes how these effects can be identified and eliminated and also describes at which locations along the beam line they cause the most problems

Mechanical vibrations. This is a particularly acute problem for horizontally-mounted microprobes compared with vertically-mounted FIBs and electron microscopes. An undamped turbo-molecular vacuum pump tends to vibrate along its vertical axis. This is not a major limitation when it is used to pump the target chamber of a vertically-mounted microscope, but is a serious consideration for a horizontally-mounted microscope because it results in a displacement of the sample relative to the focused beam position. A diffusion pump is also a poor choice for a microprobe chamber because of vibrations from the boiling oil and hydrocarbon contamination. Better choices are a magnetically-levitated turbomolecular pump in conjunction with a stiff vibration isolation unit, or an ion pump if sufficient care is taken to shield the chamber from its magnetic field.

Slit scattering. To achieve a 30 nm beam spot size with a large demagnification of 200 requires an object aperture of only 6 μm. For such small apertures a large fraction of transmitted ions pass close to the edges. Many ions are scattered from the slit edges through a very small angle because of their high momentum, and they are still transmitted along the beam line with a slightly lower energy [43]. This results in a beam halo extending many micrometers because of the chromatic aberrations of the lenses. Beam-induced hydrocarbon buildup on the aperture edges makes this problem worse, with more ions being scattered, but still transmitted through the additional deposited layer. Even after only a few days' usage of new slit edges, it has been found that this can degrade beam resolution. An additional problem is beam-induced heating, which can cause the object aperture to contract and reduce the transmitted beam current. Heavy ions result in more slit scattering and slit damage because of their greater stopping power, displacement damage, and beam heating. The easiest solution to forming a very small object aperture is to use independently adjustable slits mounted on differential micrometers [44, 45],

with each slit edge defined by a highly-polished metal cylinder, made from tungsten carbide, for example. When one portion of the cylinder is damaged, it can easily be rotated to expose a new surface to the beam, and so greatly extend the useful lifetime of the polished cylinder.

References

1. J. Goldstein, D.E. Newbury, D.C. Joy, C.E. Lyman, P. Echlin, E. Lifshin, L.C. Sawyer, J.R. Michael, Scanning Electron Microscopy and X-Ray Microanalysis, Third Edition, Plenum Press, New York, 2003.
2. P.B. Hirsch, A. Howie, R. Nicholson, D.W. Pashley, M.J. Whelan, Electron Microscopy of Thin Crystals, Krieger, Malabar, FL, 1977.
3. O.L. Krivanek, P.D. Nellist, N. Dellby, M.F. Murfitt, Z. Szilagyi, *Ultramicroscopy* **96** (2003) 229.
4. A.B. El Kareh and J.C.J. El Kareh, Electron Beams, Lenses and Optics, Vol. II, Academic Press, New York, 1970.
5. A. Septier, ed., Applied Charged Particle Optics, Advances in Electronics and Electron Physics, Supplement 13A, Academic Press, London, 1980.
6. M. Szilagyi, Electron and Ion Optics, Plenum Press, New York, 1988.
7. H. Wollnik, Optics of Charged Particles, Academic Press, London, 1987.
8. J. Orloff, ed., *Handbook of Charged Particle Optics*, CRC Press, Boca Raton, 1997.
9. L.A. Giannuzzi, F.A. Stevie, eds., Introduction to Focused Ion Beams, Instrumentation, Theory, Techniques and Practice, Springer, New York, 2005.
10. J. Orloff, M. Utlaut, L. Swanson, High Resolution Focused Ion Beams: FIB and Its Applications, Kluwer Academic/Plenum, New York, 2003.
11. P.W. Hawkes, Quadrupole Optics, Springer Tracts in Modern Optics 42, Springer, Berlin, 1966.
12. P.W. Hawkes, Quadrupoles in Electron Lens Design, Academic Press, New York, 1970.
13. K.G. Steffen, High Energy Beam Optics, Monographs and Texts in Physics and Astronomy, 17 Interscience, 1965.
14. P. Grivet, Electron Optics, Pergamon Press, London, 1965.
15. A.P. Banford, The Transport of Charged Particle Beams, E. & F. N. Spon, London, 1966.
16. A. Septier, ed., Focusing of Charged Particles, Vols. I & II, Academic Press, London, 1967.
17. G.W. Grime and F. Watt, Beam Optics of Quadrupole Probe Forming Systems, Adam Hilger, Bristol, 1984.
18. F. Watt and G.W. Grime, eds., Principles and Applications of High Energy Ion Microbeams, Adam Hilger, Bristol, 1987.
19. M.B.H. Breese, D.N. Jamieson, P.J.C. King, Materials Analysis using a Nuclear Microprobe, Wiley, New York, 1996.
20. G.W. Grime, M. Dawson, M. Marsh, I.C. McArthur, F. Watt, *Nucl. Instrum. Methods. B* **52** (1991) 52.
21. D.N. Jamieson, G.W Grime, F. Watt, *Nucl. Instrum. Methods. B* **40/41** (1989) 669.
22. M.B.H. Breese, G.W Grime, W. Linford and M. Harold, *Nucl. Instrum. Methods. B* **158** (1999) 48.
23. F. Watt, J.A. van Kan, I. Rajta, A.A. Bettiol, T.F. Choo, M.B.H. Breese and T. Osipowicz, *Nucl. Instrum. Methods. B* **210** (2003) 14.
24. J.A. van Kan, A.A. Bettiol and F. Watt, *Nano Letters* **6** (2006) 579.
25. C.G. Ryan, D.N. Jamieson, *Nucl. Instrum. Methods. B* **158** (1999) 97.
26. D. Spemann, T. Reinart, J. Vogt, D. Dobrev, T. Butz, *Nucl. Instrum. Methods. B* **190** (2002) 312.
27. R. Siegele, D.D. Cohen, N. Dytlewski, *Nucl. Instrum. Methods. B* **158** (1999) 31.

28. F.W. Martin, *Nucl. Instrum. Methods. B* **30** (1988) 242.
29. F.W. Martin, *Nucl. Instrum. Methods. B* **54** (1991) 17.
30. M.B.H. Breese, D.N. Jamieson and J.A. Cookson, *Nucl. Instrum. Methods. B* **47** (1990) 443.
31. M.B.H. Breese, D.N. Jamieson and J.A. Cookson, *Nucl. Instrum. Methods. B* **54** (1991) 28.
32. M.B.H. Breese, D.N. Jamieson,*Nucl. Instrum. Methods. B* **83** (1993) 394.
33. A. Shariff, V. Auzelyte, M. Elfman, P. Kristiansson, K. Malmqvist, C. Nilsson, J. Pallon, M. Wegden, *Nucl. Instrum. Methods. B* **231** (2005) 1.
34. M.J.J. van den Putte, J.F.J. van den Brand, D.N. Jamieson, B. Rout, R. Szymanski, *Nucl. Instrum. Methods. B* **210** (2003) 21.
35. S. Incerti, R.W. Smith, M. Merchant, G.W. Grime, F. Meot, L. Serani, Ph. Moretto, C. Touzeau, Ph. Barberet, C. Habchi, D.T. Nguyen, *Nucl. Instrum. Methods. B* **231** (2005) 76.
36. G. Datzmann, G. Dollinger, C. Goeden, A. Hauptner, H.-J. Korner, P. Reichart, O. Schmelmer, *Nucl. Instrum. Methods. B* **181** (2001) 20.
37. G. Dollinger, G. Datzmann, A. Hauptner, R. Hertenberger, H.-J.Körner, P. Reichart, B. Volckaerts, *Nucl. Instrum. Methods. B* **210** (2003) 6.
38. A. Stephan, J. Meijer, U. Weidenmuller, H. Rocken, H.H. Bukow, M. Burchard, A. Zaitsev, B. Volland, I. W. Rangelow, *Nucl. Instrum. Methods. B* **181** (2001) 39.
39. D.J.W Mous, R.G. Haitsma, T. Butz, R.-H. Flagmeyer, D. Lehmann, J. Vogt, *Nucl. Instrum. Methods. B* **130** (1997) 31.
40. J. Visser, D.J.W. Mous, A. Gottdang, R.G. Haitsma, *Nucl. Instrum. Methods. B* **231** (2005) 32.
41. E.J. Teo, M.B.H. Breese, A.A. Bettiol, F. Watt L.C. Alves, *J. Vac. Sci. and Technol. B* **22** (2004) 560.
42. F. Watt, T. Osipowicz, T.F. Choo, I. Orlic, S.M. Tang, *Nucl. Instrum. Methods. B* **136–138** (1998) 313.
43. R. Nobiling, Y. Civelekoglu, B. Povh, D. Schwalm and K. Traxel, *Nucl. Instrum. Methods.* **130** (1975) 325.
44. B.E. Fischer, *Nucl. Instrum. Methods B* **10–11** (1985) 693.
45. B.E. Fischer, *Nucl. Instrum. Methods. B* **30** (1988) 284.

Ion Spectrometers and Detectors

Tessica D.M. Weijers-Dall

The plethora of powerful research and fabrication techniques for materials characterization and processing on the nanometer scale rely critically on an equally numerous array of dedicated detectors. Two broad categories of detector can be distinguished: spectrometers and ion detectors. Spectrometers use electrostatic and/or magnetic forces to disperse or sort the incident radiation according to one of its properties, such as mass or energy before detection. Ion detectors include calorimeters as well as detectors that provide other information, such as time or position, without presorting of the incident radiation. Today, ion detectors are often incorporated into spectrometer systems forming complex, multifunctional detection systems that can make high resolution measurements of several properties of the incident radiation simultaneously.

This chapter provides an overview, from a user's perspective, of ion detectors and spectrometers commonly constituting, or incorporated within, the complex detection systems employed for nanoscale analysis with ion beams.

1 Ion Detectors

Charged particles are not detected directly, but are instead detected based on their interaction with the material of the detector. Passage of the particle through the detector causes it to expend its energy in the form of electronic energy deposition as well as during events such as nuclear collisions (see chapter "Basis of Ion Scattering in Nanoscale Materials"). For determination of information such as time and position, what is required of the detector is that it can, often with high efficiency and resolution, detect the presence of the incident radiation. For calorimetry, it is important that the incident particle, and its secondary emissions, expend all of their energy within the active volume of the detector (since the total energy is determined from the integral of the stopping force).

T.D.M. Weijers-Dall (✉)
Research School of Physical Sciences and Engineering, The Australian National University, Canberra ACT 0200, Australia
e-mail: tessica.dall@aihw.gov.au

1.1 Electron and Photo-Multipliers

Electron and photo-multipliers find application for direct ion detection, but are also commonly found in systems detecting secondary emission of electrons or photons after the passage of an ion, for example through the start or stop foil in a time-of-flight (TOF) system. Their very high gain characteristics make them ideally suited to single ion detection. Arrays of electron multipliers (multichannel plates) find common application in applications where position sensitivity or timing is important.

1.1.1 Electron Multipliers, Channeltrons and Multichannel Plates

Electron multipliers, or discrete dynode electron multipliers to be precise, consist of a series of dynodes maintained at increasingly positive potentials (see Fig. 1(a)), up to several thousand volts. The electric field between these dynodes serves to accelerate the emitted electrons on to the next stage. The impact of each accelerated electron on the dynode results in the ejection of approximately 2–4 lower energy electrons. Over several stages this results in an avalanche of electron emission, initiated by the impact of a single quanta of incident radiation on the first dynode, giving typical amplification factors of several million. For this reason, electron multipliers are also referred to as avalanching ion detectors. Finally, the arrival of the electrons at the anode results in a sharp spike of current at the output of the detector.
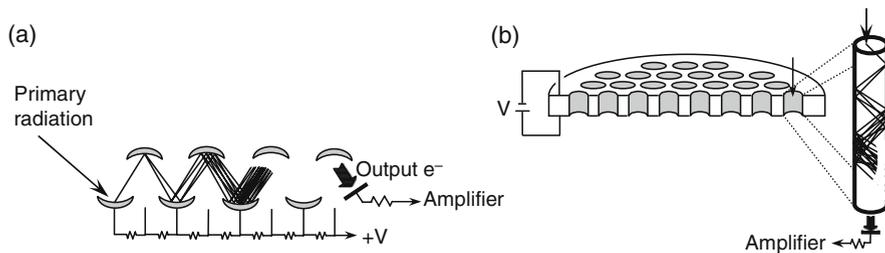
Fig. 1 (**a**) Principle of operation of an electron multiplier illustrated in the discrete dynode case, and (**b**) schematic of a microchannel plate

Channeltrons operate along the same principles, except that the discrete dynode structure is replaced by a curved, horn-shaped, continuous dynode. This dynode is coated with a high resistance material which ensures a voltage gradient is produced along the length of the detector. Repeated collisions of the electrons with the walls of the channeltron produce an avalanche of electrons in much the same way as the discrete dynode system. Channeltrons are used as secondary electron detectors in proton beam writing and nuclear microscopy (see chapter "Proton Beam Writing: A New 3D Nano Lithographic Technique").

Two-dimensional arrays of miniature parallel-walled continuous dynode electron multipliers, called micro-channel plates or MCP's (see Fig. 1(b)) find widespread application. MCP's are employed in mass spectrometry where the spatial distribution of the incident ions provides information about the ions mass, and are also used as position sensitive detectors in the small magnetic spectrometers used for medium energy ion scattering (see chapter "Low and Medium Energy Ion Scattering for Near Surface Structure and Nanoscale Depth Profiling"). Their small channel diameter to length ratio also provides excellent timing properties, hence their application as time detectors in ToF-ERDA (see chapter "Thin Film Characterisation using MeV Ion Beams").

Electron multipliers in all their forms provide the advantages of very high gains/sensitivity, high output currents (μA), and fast response times. They are able to detect both positively and negatively charged ions. Electron multipliers are typically made of a lead silicate glass and are therefore very fragile. Many of the dynode coating materials are also hydroscopic, which means that their lifetimes are often short and limited by the number of exposures to atmosphere. Ceramic and glass-ceramic versions are generally considered more robust for applications which require frequent exposure to air. MCP's in particular are sensitive to moisture absorption, which can cause buckling and cracking. However, there is intense development in this area and newer types are less susceptible.

1.1.2 Photomultiplier Tubes and Scintillation Detectors

Photomultipliers combine an electron multiplier with a photocathode. Incident photons strike the photocathode and eject electrons through the photoelectric effect. These are then avalanche-multiplied to form a detectable signal by the electron multiplier. Photomultipliers are also commonly referred to as photomultiplier tubes or PMT's, since the whole structure is encased in an evacuated glass tube. Photons pass easily into the detector, however the sensitive elements of the detector are shielded from contamination and moisture, which significantly increases the lifetime of the detector.

PMT's also find limited application in scintillation detectors, which couple a scintillation medium with the PMT. This medium fluoresces when struck by ionising radiation, with different scintillators providing different operational characteristics. Scintillation detectors find application as gamma ray detectors in resonant nuclear reaction analysis. More recently, photo-diodes are taking over as readout devices because of their higher quantum efficiency and small size. This enables them to be used as readout devices for X-ray imaging applications.

1.2 Solid State Detectors

One of most commonly used detectors is the solid state detector (SSD). SSD's are reliable, compact, and simple to use and operate. They are most often made of either

silicon or germanium, and employ a fully depleted, reverse-biased diode structure
(see Fig. 2). Electron-hole pairs created by the passage of the incident particle are
separated by the strong electric field created by this external bias. Electrons drift
toward the anode, holes to the cathode, producing a current pulse on the electrode.
A charge-sensitive pre-amplifier can be used to collect this charge. The time integral
of the current pulse is proportional to the total energy deposited by the particle in
the detector.
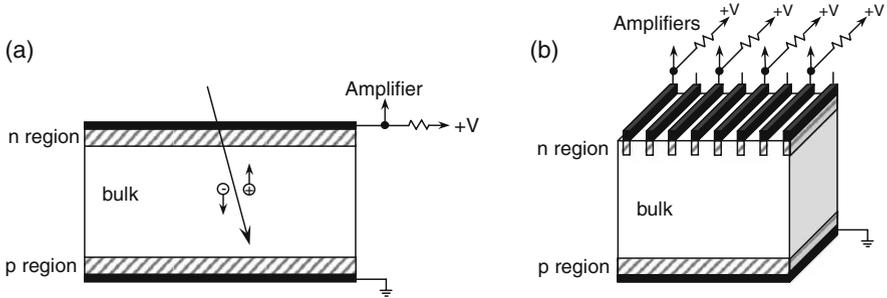
Fig. 2 (**a**) Schematic of a solid state detector, and (**b**) cross-section through a silicon strip detector

The number of electron-hole pairs, N, produced in the detector is related to the
electronic energy loss of the incident particle, E:

$$N = \omega E$$

where ω is the average electronic energy deposition required, at a given temperature,
to produce a single electron-hole pair.

For semiconductors, ω is essentially independent of the energy and type of ionis-
ing radiation. For this reason, SSD's are very versatile detectors. ω is also essentially
constant in gaseous detectors. However, it is much larger in this case e.g. 26 eV for
an electron-ion pair in Ar compared to 3.6 eV for an electron-hole pair in Si. For
the same incident radiation therefore, many more electrons are produced in solid
compared to gaseous detector materials, resulting in much larger pulses. Since the
main source of noise in the detector itself is the statistical fluctuations in the number
of charge carriers produced, this means that solid state detectors provide intrinsically
higher energy resolution, particularly for light ions such as alpha's and protons (13–
18 keV FWHM for alphas in a Si surface barrier detector [1] and 9 keV FWHM for
protons). SSD's also provide high spatial resolution (5–10 μm [1]). Furthermore, the
Si bandgap (1.12 eV) is quite large compared to room temperature thermal energies
(kT = 1/40 eV), providing excellent signal to noise ratios and obviating the need for
cooling except in ultra-low noise applications.

The energy resolution of SSD's for heavier ions is much poorer. However, it is
primarily their susceptibility to radiation damage that impedes the use of SSD's for

heavy ion detection. Cooling of the detector may help in this regard, but in general radiation damage greatly reduces the lifetime of a SSD.

It is also recognised that, particularly for heavier ions, a small variability of ω is observed. Most commonly this results in lower than expected pulse heights with increasing atomic number and energy, hence this effect is often referred to as the pulse height defect or deficit effect. The origins of these effects are not entirely known, however, over small energy ranges and for single species applications the variability of ω is often negligible. For high precision applications several compensation and calibration schemes have been proposed ([2–5] amongst others).

There are several types of SSD that can be employed. Surface barrier and PIN detectors are common, whereas silicon strip detectors and charge-coupled devices (CCD's) incorporate position sensitivity. Species sensitivity can be achieved by incorporation of the SSD into a larger detection system. For example a high resolution time-of-flight system to provide momentum information or through the use of ultra-thin Si detectors in a compact $\Delta E - E$ telescope where the initial rate of ionisation energy loss is used to determine atomic number.

SSD's require only a moderate vacuum ($< 1 \, \text{kPa}$) for operation and the biasses required for full depletion are only of the order of 20–100 V depending on the detector type and thickness. The optimum bias generally increases with time in heavy ion applications due to radiation damage and should be optimized to balance signal to noise levels and pulse resolution. In high resolution applications consideration should also be given to the thickness and composition of the passivation layer, or contact layer in the case of a surface barrier detector, having regard for the energy loss as well as the energy straggling of the ion in the layer.

1.3 Ionisation Chambers

One of the oldest and most versatile types of ion detector is the ionisation chamber. While SSD's often provide higher energy resolution, ionisation chambers due to their larger physical size are particularly suited to large area and large solid angle applications. Ionisation chambers are often incorporated into larger detection systems where they provide ΔE or residual energy measurement but their strengths are best exploited in applications where multiple properties of the ion, including energy, atomic number and position, are of simultaneous interest. Ionisation detector design and operation is more complex than for their solid state counterparts, but they have the distinct advantage of direct and readily incorporated species identification.

In an ionisation chamber, an electric field generated by two electrodes within a gas-filled volume, is used to separate electron-ion pairs formed along an incident ion's path in the detector. The magnitude of the electric field is critical. Most commonly, chambers used for ion detection are operated in the "ionisation" or "saturation" regime (see Fig. 3(a)), where the number of electrons produced by ionization

is proportional to the amount of energy deposited by the ion and is independent of changes in the applied potential (within the saturation region). Ionisation electrons drift toward the anode creating a measurable charge pulse or current. The physical properties of the ion can be determined by measuring the magnitude, temporal and spatial distribution of this ionisation within the volume of the detector.
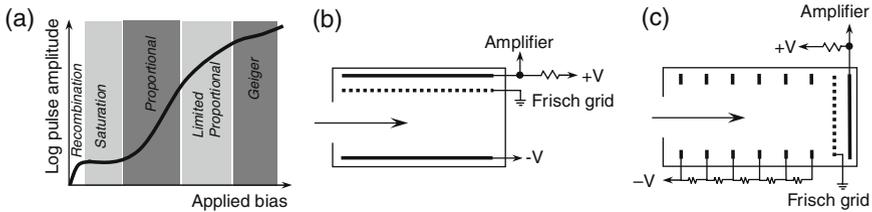
Fig. 3 (a) Pulse amplitude as a function of applied bias showing different operation regimes of a gas chamber, (b) parallel plate gas detector geometry, and (c) Bragg detector geometry

Ionisation detectors typically use either a parallel plate or Bragg electrode geometry (see Fig. 3(b) and (c)). The primary difference is the direction of the electric field. In the Bragg geometry the electric field is parallel to the ion path, hence the charge collected on the anode as a function of time will reflect the spatial distribution of charge deposited in the detector. Since this distribution follows the electronic energy loss curve and includes the characteristic Bragg maximum, the Bragg geometry intrinsically incorporates species resolution.

In the parallel plate geometry, the electric field is perpendicular to the ion path. Hence species identification is achieved by dividing the anode into sections and collecting electrons over the first portion of the detector, where the rate of ionization is proportional to the atomic number of the ion. Furthermore, 2D position information can also be achieved with little added complexity, by dividing a section of the anode into backgammon shaped right and left halves to provide lateral spatial information from the relative size of the current pulse from the two halves and utilizing the cathode signal to provide an entrance height determination.

In both geometries a grounded Frisch grid is employed. The purpose of this grid is to shield the anode so that it is sensitive only to electrons once they pass into the area between the grid and the anode. This ensures that the spatial distribution of the ionization cloud left by the ion is not lost, since the drift time of the electrons to the anode is much longer than the time of passage of the ion in the detector.

Most often the ionisation medium of choice is a gas, although both liquid and solid media are also possible. Noble gases such as Ar, Kr and Xe are common, as are hydrocarbons such as methane and propane. A gas flow system can be used to refresh the detector gas every few minutes preventing pulse height degradation over time. This makes the detector largely immune from heavy ion radiation damage and making them well suited to heavy ion detection. Pressures up to several hundred mbar are achievable, reducing the physical size of the detector, but this must balanced against the thickness of the window used to separate the gas in the detector

from the vacuum in the chamber since considerable energy straggling and energy loss is introduced by the use of large window thicknesses.

Like their solid state counterparts, it has recently been found that the response of ionisation chambers shows a slight species and energy dependence, that is ω is non-constant, particularly for heavy ions. Very precise energy spectroscopy therefore requires the application of a (potentially gas-specific) energy correction to account for this (see [6] for an example of such a scheme).

Gas ionisation detectors can achieve energy resolutions better than 1% [7–9] for heavy ions exceeding that possible with solid state detectors, which provide excellent energy resolution for light ions. Due to electron losses resulting from field inhomogenities the Bragg geometry is generally considered to provide better energy resolution, although field shaping can be incorporated into the parallel plate design.

2 Spectrometers

Spectrometers spatially disperse an ion beam according to one of its properties in much the same way as a prism is used to sort light according to its wavelengths. This is often desirable in high resolution measurements, but it can also be used to increase the efficiency of detection or to allow weaker signals to be isolated from stronger carrier beams. Both magnetic and electrostatic spectrometers are employed for this purpose.

A spectrometer must be combined with a detection system, which may be as simple as a particle counter or comprise a sophisticated ion detector that provides additional information on the particles of interest. Both small and large area ion detectors find application in spectrometers. Small area detectors must either be scanned across the output of the spectrometer, a variable field strength used to scan the output beam across the detector face, or used to form an array of detectors across the field of interest.

2.1 Magnetic Spectrometers

Magnetic spectrometers work on the principle that the trajectory of a charged particle in a magnetic field is determined by its magnetic rigidity. Particles with varying magnetic rigidities, given by the ratio of the particle's momentum to charge, are therefore spatial dispersed across the focal plane of the spectrometer enabling high resolution detection (see Fig. 4(a)).

Magnetic spectrographs can be operated in two modes depending on whether the magnetic field region is a vacuum or filled with gas. The vacuum mode is often employed in high resolution particle spectroscopy applications, while the gas-filled mode is particularly useful for applications, such as heavy ion reaction recoil spectroscopy, where the particles of interest exhibit a charge state distribution. Under vacuum, particles of the same energy but different charge states will undertake
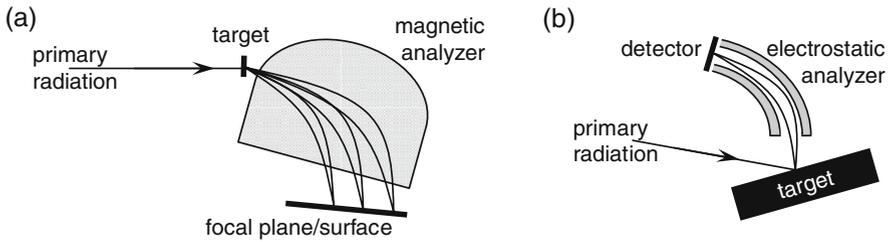
Fig. 4 Schematic illustration and principle of operation of (**a**) a magnetic spectrometer and (**b**) an electrostatic spectrometer

separate trajectories since their magnetic rigidities differ. If the field region is filled with gas, however, the particles will experience charge-exchanging collisions with the gas atoms such that their resultant trajectories will merge to approximately follow that of the mean charge state.

There are a large variety of magnetic spectrometer designs. Two of the most common are the split-pole or Enge and the Q3D. The split-pole design is characterized by having two separate pole pieces magnetized by the same set of coils. This design enables good correction for second order aberrations and is stigmatically focusing over the whole range. The term Q3D refers to a series of spectrometer designs incorporating one quadrupole and 3 dipole elements. In a Q3D spectrometer a (transverse) image is formed in the middle of the spectrometer which is then projected to form another image at the output focal plane. This enables relatively simple correction for both second and higher order aberrations. As a consequence, Q3D spectrometers exhibit resolving powers exceeding 10^4 for solid angles of 10–15 msr [10]. The split pole design also has a large acceptance solid angle of 8–12 msr, however, higher order aberrations limit the effective solid angle to approximately 2 msr.

Magnetic spectrometers are large-scale devices with dispersion distances measured in meters.

2.2 Electrostatic Spectrometers

Electrostatic spectrometers or analyzers (ESA's) are filter-like devices consisting of two concentric and parallel curved metallic plates. These plates are held at different potentials such that they deflect the path of incident particles (see Fig. 4(b)). Only particles with energies in a narrow band are able to pass, the remainder are lost by impinging on the walls of the analyzer. Particles exiting the analyzer are collected and detected. Energy spectra can be obtained by scanning the potential on the plates in a step-wise fashion. Similarly to magnetic spectrometers, ESA's are charge state sensitive.

Electrostatic analyzers are employed for electron and low energy ion detection and are well suited to surface analysis applications. They are compact in design, and can be fitted into a standard vacuum chamber, but typically have small acceptance solid angles. The angular acceptance angle is typically 0.15 msr, but varies with the size of the entrance aperture of the analyzer. Higher efficiency designs include an ion lens system to collect ions/electrons from a larger solid angle. ESA's are well regarded for their excellent energy resolution. $\Delta E/E$ values of 10^4 are commonly reported.

References

1. J.R. Tesmer and M. Nastasi, editors: *Handbook of Modern Ion Beam Analysis*. Materials Research Society, Pittsburgh, 1995
2. H.W. Schmitt, W.E. Kiker, and C.W. Williams: Phys. Rev. B **137**, 837 (1965)
3. B.D. Wilkins, M.J. Fluss, S.B. Kaufman, C.E. Gross, and E.P. Steinberg: Nucl. Instrum. Methods **92**, 381 (1971)
4. E.P. Steinberg, S.B. Kaufman, B.D. Wilkins, C.E. Gross, and M.J. Fluss: Nucl. Instrum. Methods **99** 309 (1972)
5. S.B. Kaufman, E.P. Steinberg, B.D. Wilkins, J. Unik, A.J. Gorski, and M.J. Fluss: Nucl. Instrum. Methods **115**, 47 (1974)
6. T.D.M. Weijers, T.R. Ophel, H. Timmers, and R.G. Elliman: Nucl. Instrum. Methods A **483**, 676–688 (2002)
7. W. Assmann: Nucl. Instrum. Methods B **64**, 267 (1992)
8. A. Oed, P. Geltenbort, F. Goennenwein, T. Manning, D. Souque: Nucl. Instrum. Methods **205**, 455 (1983)
9. T.R. Ophel, H. Timmers, and R.G. Elliman: Nucl. Instrum. Methods A **423**, 381 (1999)
10. H.A. Enge: Nucl. Instrum. Methods **162**, 161 (1979)

Electronics for Application of Ion Beams in Nanoscience

Harry J. Whitlow

1 Introduction

The breadth of ion-beam nanoscience methods discussed in the proceeding chapters implies that the control and readout systems also span a wide range of techniques. For many nanoscience and technology applications the need to process or characterise a batches of samples has lead to a convergence of data acquisition systems and control systems. This coupled with the development of high-performance low-cost computing over the past decades has lead to a progressive evolution in data acquisition technology. Historically, much of the electronics used in nanoscience and technology using ion beams is based on nuclear analogue pulse counting and spectrometry where NIM (Nuclear Instrumentation Modules) have been a standard tool for some decades. Typically, the pulse amplitude and time of occurrence is measured, e.g. in Medium Energy Ion Scattering (MEIS) (chapter "Low and Medium Energy Ion Scattering for Near Surface Structure and Nanoscale Depth Profiling") and Time of Flight-Energy Elastic Recoil Detection Analysis (ToF-E ERDA) (chapter "Thin Film Characterisation by means of MeV Ion Beams"). Figure 1 illustrates a typical pulse spectroscopy electronics chain. A comprehensive overview of pulse spectroscopy practice is given in Leo's classic text [1]. A preamplifier converts the weak pulse of charge from the detector to a voltage signal which is transmitted to signal processor that either produces a signal proportional to the amplitude of the voltage pulse, or the time difference between the pulse and some reference time signal. The output signal is then digitised using an *analogue to digital converter* (ADC) so the spectra can be displayed and stored in a data acquisition system. This is often based on NIM modules in a crate. However, there is a trend towards compact and lower cost devices based on application specific integrated circuits (ASIC) where the entire spectroscopy chain (Fig. 1) is integrated into a single unit.

H. J. Whitlow (✉)
Department of Physics, University of Jyväskylä, FIN-40014, Jyväskylä, Finland
e-mail: harry.j.whitlow@jyu.fi
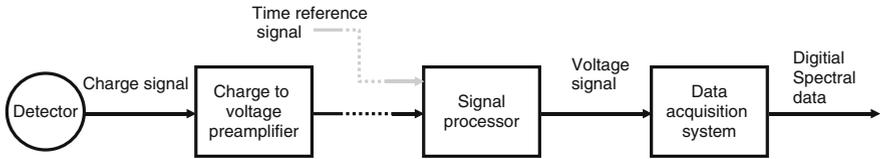
Fig. 1 Schematic illustration of a conventional pulse spectroscopy electronics chain. This is often based on NIM modules in a crate. The signal processor can be a spectroscopy amplifier for measurement of amplitude spectra and time-to-amplitude converter for measurement of time spectra

2 Analogue and Digital Signal Handling

The charge pulse from the detector is converted to a voltage pulse in a *charge-sensitive preamplifier* (shown schematically in Fig. 2) using an operational amplifier circuit configured as a Miller integrator which integrates the current from the detector. The output from the preamplifier is a series of voltage steps with amplitude that corresponds to the size of the charge pulse flowing from the detector. A metal oxide semiconductor (MOS) transistor is used as the input stage of the amplifier to achieve high sensitivity for the extremely small current pulses (i_{det}) from the detector. The conversion gain (Volts/Coulomb) is governed by the value of the feedback capacitor C_f. This capacitor is usually slowly discharged by a high value resistor in parallel which gives a long exponential decaying tail to the voltage step. Alternatively a second MOS transistor in parallel with C_f is used to restore the integrator output to
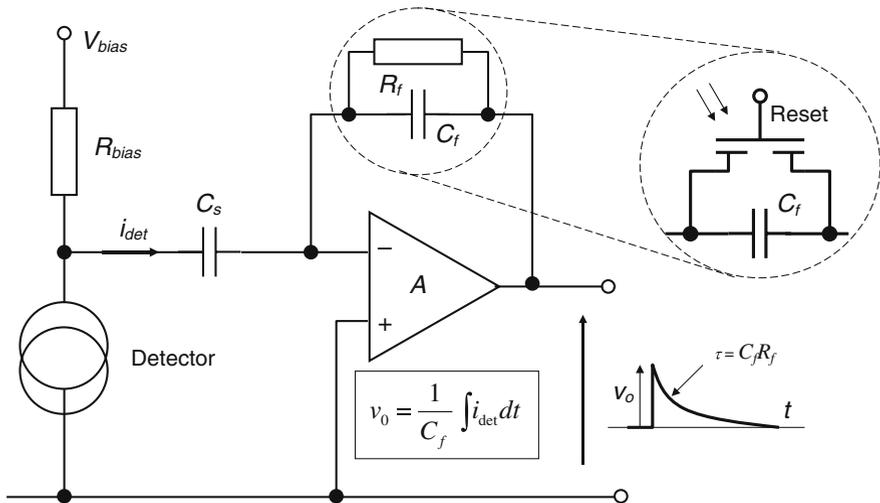
Fig. 2 Schematic illustration of a generic charge-sensitive preamplifier. A DC bias voltage can be applied to the detector via a bias resistor, R_{bias}, chosen to be large so the current pulse flows to the operational amplifier. The resistor R_f which slowly discharges to feedback capacitor to restore the output level voltage v_o to zero, may be replaced by a MOS transistor (inset) that discharges the capacitor either by biasing the gate or by illuminating the gate with a light pulse

zero. Applying a bias pulse to the gate or illuminating it with a pulse of light resets the preamplifier output to zero. Miller integrator circuits are widely used for other applications in ion beam technology. For example, *charge integrator* modules are used to measure the number of ions impinging on a target by integrating the electrical current from the target (using suitable compensation for secondary electrons). Time to amplitude converters (TAC) (discussed below) are also based on Miller integrator circuits.

In nanoscience and technology one is often interested in measuring extremely small changes in signals. This requires high resolution implying the electronic noise performance of the front-end preamplifier is crucial. The electronic noise is conveniently expressed in terms of the *electronic noise charge* (ENC) in units of e^- which is related to the FWHM energy resolution in eV, $\Delta E_{FWHM} = 2.35\eta ENC_{RMS}$, where η is the deposited energy required to create an electron-pair. Generally the ENC is dominated by the thermal noise in the MOS input transistor [3]:

$$ENC_{RMS} = \frac{2.71\,(C_{det} + C_{cable} + C_{MOS})}{e}\sqrt{\frac{\Gamma k T}{3 g_m \tau_p}} \tag{1}$$

Equation (1) illustrates how the ENC can be optimised. The excess noise factor Γ is fixed at 1.5–1.8. Minimising the detector, cable and MOS transistor input capacitances, (C_{det}, C_{cable} and C_{MOS}, respectively) have the greatest effect because these terms are outside the square root. It can be shown that for minimum electronic noise the detector capacitance should be, $C_{det} + C_{cable} = 3C_{MOS}$. The MOS input transistor and its operating point should be chosen so the transconductance g_m is large and at the same time the gate-source capacitance, C_{MOS} matches the detector. Cooling the input transistor to reduce, T, the absolute temperature, will reduce the ENC. Cooling to L-N$_2$ temperatures reduces the ENC by about a factor 2 compared to room temperature. The choice of the peaking-time τ_p governs the time period over which the signal is integrated over by the electronics chain [2]. Larger τ_p values give a better amplitude resolution by averaging out the ENC at the expense of degrading the time resolution. The upper limit for τ_p is governed by the onset of low-frequency $1/f$ dependent flicker noise. Typically, τ_p is chosen to be a few μs to obtain best pulse amplitude resolution and 10–100 ns if time information is more important.

In addition to electronic noise, the resolution can be significantly degraded by pick-up of spurious signals originating from so called earth- or ground-loops or introduced by inductive and capacitive coupling [3]. Meticulous attention must be paid to grounding practice so as to avoid ground loops. This generally requires: (i) Separated measurement and machine power cabling where the earth-lines are joined as close to a low-impedance building-earth as possible. (ii) Electrically insulating vacuum tube sections to electrically isolate sample chambers from the beam-line and pumps. (iii) Connection of detectors so that the grounded shielding that is connected to the signal ground is electrically isolated from the vacuum chamber. (iv) The signal ground should be connected as close to the detector element as possible. This last point is especially important for fast timing electronics, e.g. in Tof-ERDA (chapter

"Thin Film Characterisation by Means of MeV Ion Beams"), because even a cm or two of low resistance solid metal can present a high impedance to signals whose power spectrum is in the GHz region. To minimise inductive and capacitive pick-up, the preamplifier input line should be connected to the detector with shielded, or tight twisted-pair cables that are as short as possible. This also minimises the contribution of C_{cable} to the ENC (Eq. 1). In this connection fast transient signals, such as negative NIM logic signals should also be well shielded and physically separated from the input signal lines to prevent cross-talk. Another source of cross-talk is power-supply coupling. The use of common DC supply lines for pre- and main-amplifiers as well as fast logic should be avoided because this can introduce disturbing signals such as switching transients into the preamplifier circuits.

The voltage step signal from the preamplifier is generally too small to be used directly. A *spectroscopy amplifier* module is used to amplify and filter the signal to average out the effect of random noise and spurious pick-up e.g. 50/60 Hz from mains power. If the pulse count rate is not so high that they overlap in time, the optimum signal to noise ratio is obtained when the output pulse has a Gaussian shape. Successive integration and differentiation of the pulse in a series of 2–4 $CR - RC$ stages is generally used in spectroscopy amplifier modules to give a semi-Gaussian output pulse. The time constant $\tau = CR$ for the $CR - RC$ shaping is generally selected between 0.5 and 10 μs for precise measurement of amplitude. Normally, to allow compensation of under- and overshoot of the semi-Gaussian pulse that leads to a baseline shift, especially at high count rates an adjustable pole-zero correction is provided. The shaping time and peaking time (Sect. 2) are related, with $\tau = 2.2\tau_p$.

In order to extract timing information from the preamplifier pulse, discriminator modules are used which provide an output pulse when the leading edge of the pulse exceeds some threshold value. A *timing filter amplifier* module ($\tau = 30 - 200$ ns) or a wideband amplifier module is often used in place of the spectroscopy amplifier to boost the detector signal before feeding it to the discriminator input. The *constant fraction discriminator* is most widely used because output pulse is produced when the pulse amplitude exceeds some fixed fraction of its maximum amplitude. This gives better timing accuracy than a *leading edge discriminator* for pulses with identical shape but different amplitudes.

In order to measure time differences, i.e. in ToF-ERDA (chapter "Thin Film Characterisation by Means of MeV Ion Beams"), a *time to amplitude converter* (TAC) is used. This is based using a charge integrator (Fig. 2) where the detector is replaced by a constant current source. These function by integrating a constant current pulse whose duration is governed by start and stop pulses so that the final amplitude is proportional to the time difference between the pulses.

ADCs are used to digitise the output signals from spectroscopy amplifiers and TACs so they can be stored and processed by a computer. For pulse spectroscopy *peak-sensing ADC* modules are used which incorporate an analogue sample and hold circuit that converts the maximum amplitude of the pulse to a constant analogue level which is digitised. Peak-sensing ADCs usually also incorporate circuits so the pulse is only converted if a digital gate pulse is applied in coincidence with the analogue pulse.

Two general modes of *data acquisition* are used for pulse spectroscopy in nanoscience and technology. The simplest is *multi-channel analyser* mode where the contents of a computer memory cell with address corresponding to the ADC conversion value is incremented. This produces a single-dimensional histogram of pulse heights stored as sequential channel contents in computer memory. An example is a backscattering spectrum. The other mode is *event mode* where data is stored as sequential event records where each event record is made up of conversion records from different ADC's etc. The drawback of this approach is the need for a trigger system that becomes increasingly complex as the number of channels increase. To alleviate this problem, a new event-mode approach is emerging, total data readout (TDR) [4] that requires only minimal trigger circuity. This is based on time-stamping the output and identity of ADC conversion values. In both TDR and event mode collection the data file is histogramed (often on-line) to produce single or multidimensional spectra.

3 Modular Electronics

The electronics readout and control systems used for control and data acquisition are based on functional modules. A number of modular standards have been defined for mechanical specification, connector types and signal levels for different standards, such as NIM, computer animated measurement and control (CAMAC) [1] and VME (VersaModular Eurocard) VME as well as PCI (Peripheral Component Interconnect). All except NIM have functionality for computer readout and control. The above mentioned legacy from experimental nuclear physics has lead to the widespread use of NIM modules for pulse spectroscopy. This is based around a $19''$ crate into which the modules slot. Power to the modules is supplied by a backplane bus structure. The NIM standard defines standard signal levels and impedances [1, 5] that allows the modules to be interconnected simply in different configurations. A large number of different modules are available including digital counters and logic units with AND, OR and NOT functions, as well as beam current integrators and pulse stretchers, delay units and calibration pulse generators. For systems where many identical readout channels are needed, such as for a high-resolution strip-line detector for a magnetic spectrometer, it may be effective and economical to use an application specific integrated circuit (ASIC). These circuits are extremely compact which facilitates incorporating the entire readout chain from preamplifiers to ADC within the detector housing. Figure 3 compares a crate of NIM electronics with an ASIC pulse spectroscopy readout chip with equivalent functionality by integrating the pre-amplifier-shaper amplifier and analogue sample and hold circuits for spectroscopy with fast trigger logic on the same chip. The drawback of ASIC technology is that the manufactured circuit configuration is fixed. However, implementation is rather straightforward because of the availability of tested circuit modules in a library as well as simulation tools that allow the circuit design to be rigorously tested before submission to silicon-foundry for fabrication.
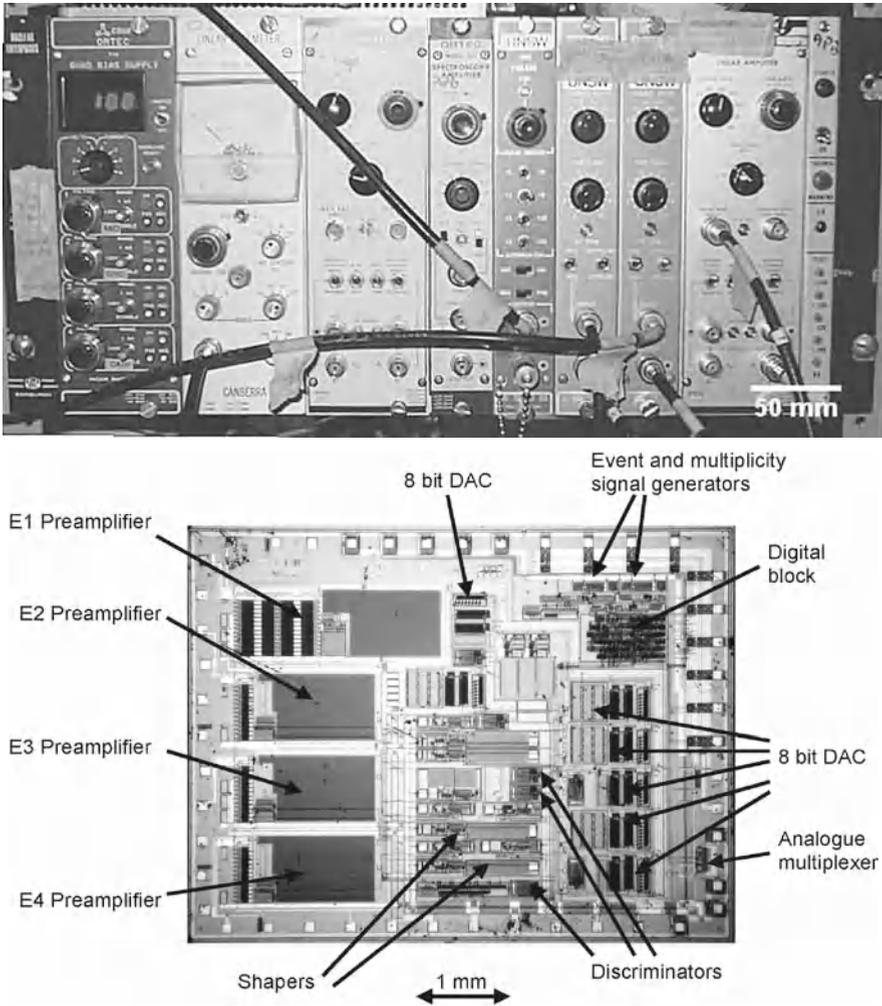
Fig. 3 *Top*: Crate of NIM electronics modules. *Bottom*: ASIC chip for readout with an equivalent functionality. This was achieved by combining low-level analogue preamplifers, spectroscopy electronics with fast digital timing on the same chip [3]

The advent of high-speed ASIC electronics with transistors with cut-off frequencies in excess of 30–100 GHz (chapter "High Speed Electronics") is leading to an evolution towards digital signal processors that employ direct digitisation and of the current pulses from the detector. This gives considerable advantages because the information in each pulse can be analysed and filtered in a mathematically optimal way, and using TDR, much higher data collection speeds can be achieved than is possible in a conventional analogue electronics system. The prospect for nanoscience and technology using ion beams is that smaller signals e.g. from a

surface layer can be measured more accurately in the presence of otherwise over-whelming signals from the substrate.

4 Control Systems

The use of ion beam methods for characterisation and processing of nanometre structures often requires computer control. Such automated control relieves the operator of tedious and repetitive tasks and improves repeatability. In their simplest form these control systems may monitor a single parameter and stop the process when the end-point is reached e.g. in plasma processing (chapter "Plasma Etching and Integration with Nanoprocessing"). At the other end of the scale a full-blown control system for a FIB instrument (chapter "Focused Ion Beam Machining and Deposition") performs controls functions such as beam rastering and blanking, controlling the the ion gun and stabilises the mass flow of the reactive gases used for etching and deposition.

Complex control systems are usually based on a hierarchy approach. At the lowest level, single controllers are used to control individual *plant* items such as a magnet power supply. *Closed-loop* controllers are generally used where plant output (e.g. magnetic field, feedstock gas flow, temperature) must be maintained at some set-point. *Open-loop* controllers are suited to simple actions or where the output values are well defined (e.g. beam on-off, change-sample, readout and storage of data from a data acquisition system etc.). At the next level, groups of controllers act together to control basic functions such as the orientation of a sample on channelling goniometer, vacuum system control, the ion source or beam rastering. The highest level of control brings together the low-and intermediate-level control functions. These can perform all the steps necessary to measure, or process an entire batch of samples without operator intervention. For example, in Medium Energy Ion Scattering (MEIS) measurements (chapter "Low and Medium Energy Ion Scattering for Near Surface Structure and Nanoscale Depth Profiling") high-level control could include positing the sample, ramping the electrostatic or magnetic field, collection of data and storing the data as computer files. Generally, the computational tasks in process and measurement control are well within the capability of a modern PC. In implementing computer control hardware lifetime and upgrading should be taken into account. PCs generally have a useful life of five years, or so whereas electronics modules and plant often have a useful life that is many times longer. For this reason, controllers should facilitate upgrading by using standard interfaces and signal standards (RS232, USB, IEEE-bus, TTL). Computer-based control systems can be programmed in a variety of standard compilers and interpretive software, such as VisualBasic, C++ and also LabView^{TM}. These allow combination of boolean logic with conventional controllers such as Proportional Integral Derivative (PID) closed-loop controllers for linear systems with fuzzy logic which is suitable for non-linear control tasks. Standard driver routines for these environments are often provided by the controller manufacturer. The choice of programming environment depends on the particular application. Where high-speed and processing a large amount of
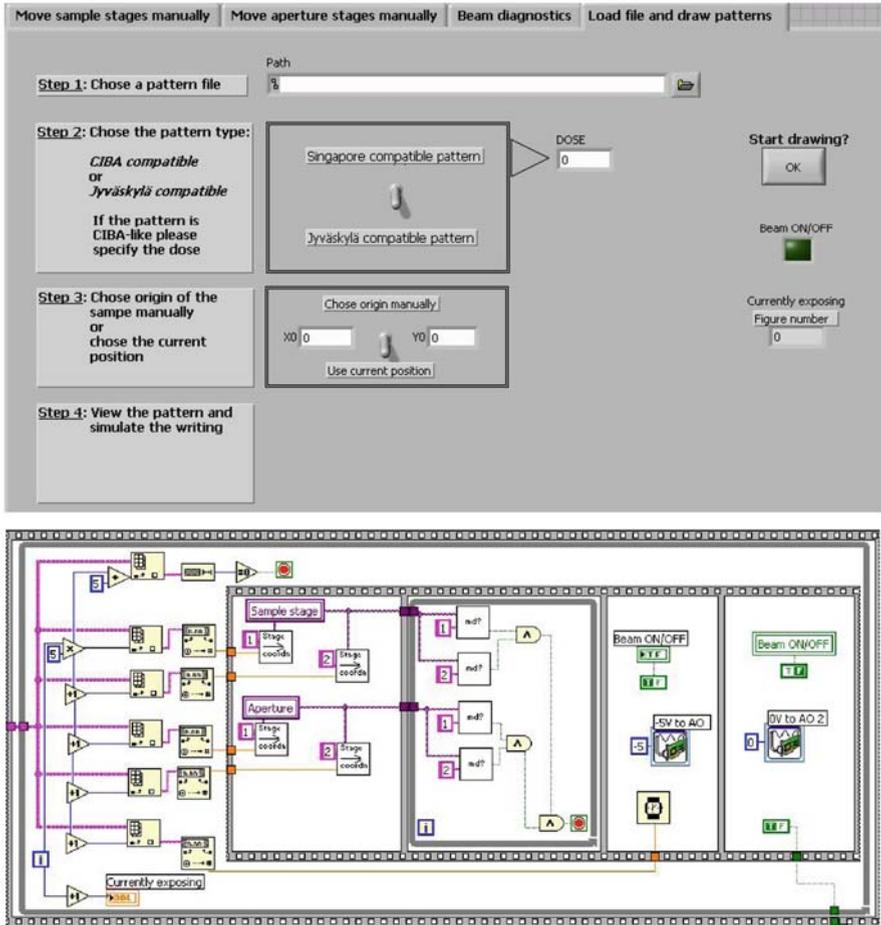
Fig. 4 Screen images of a prototype LabView™ control system for MeV ion beam lithography (chapter "Proton Beam Writing: A New 3D Nano Lithographic Technique"). *Top*: control panel for pattern writing. *Bottom*: part of the corresponding block diagram. (With kind permission from Sergey Gorelick [6])

information from a single measurement is important a fully compiled environment such as C++ is preferable. The drawback is that development, or even small changes, require a considerable personnel and time investment. For prototyping and where speed is not critical, LabView™ offers a flexible graphical programming environment that allows complex measurement and control functions to be built up in a simple way. No special programming skills are needed. An example is shown in Fig. 4 which shows screen views of a prototype control system and part of the corresponding block diagram for MeV ion beam lithography system (chapter "Proton Beam Writing: A New 3D Nano Lithographic Technique").

References

1. W. R. Leo, *Techniques for nuclear and particle physics experiments*, (Springer, Berlin, 1987).
2. E. Nygård, Ph.D. Thesis, U. Oslo Internal Report, No. 230, 1997.
3. L. Carlén, G. Førre, P. Golubev, B. Jakobsson, A. Kolozhvari, P. Marciniewski, A. Siwek, E. J. van Veldhuizen, L. Westerberg, H. J. Whitlow, and J. M. Østby, Nucl. Instrum. Methods A 516 (2004) 327–347.
4. I. H. Lazarus, D. E. Appelbe, P. A. Butler, P. J. Coleman-Smith, J. R. Cresswell, S. J. Freeman, R. D. Herzberg, I. Hibbert, D. T. Joss, S. C. Letts, R. D. Page, V. F. E. Pucknell, P. H. Regan, J. Sampson, J. Simpson, J. Thornhill, and R. Wadsworth. IEEE Trans. Nucl. Sci. 48 (2001) 567–569.
5. The Ortec Products on-line catalogue, http://www.ortec-online.com/products.htm
6. S. Gorelick, *Labview^{TM} control-software for the MeV ion beam programmable proximity aperture lithography system*, Private communication 2007.

Index