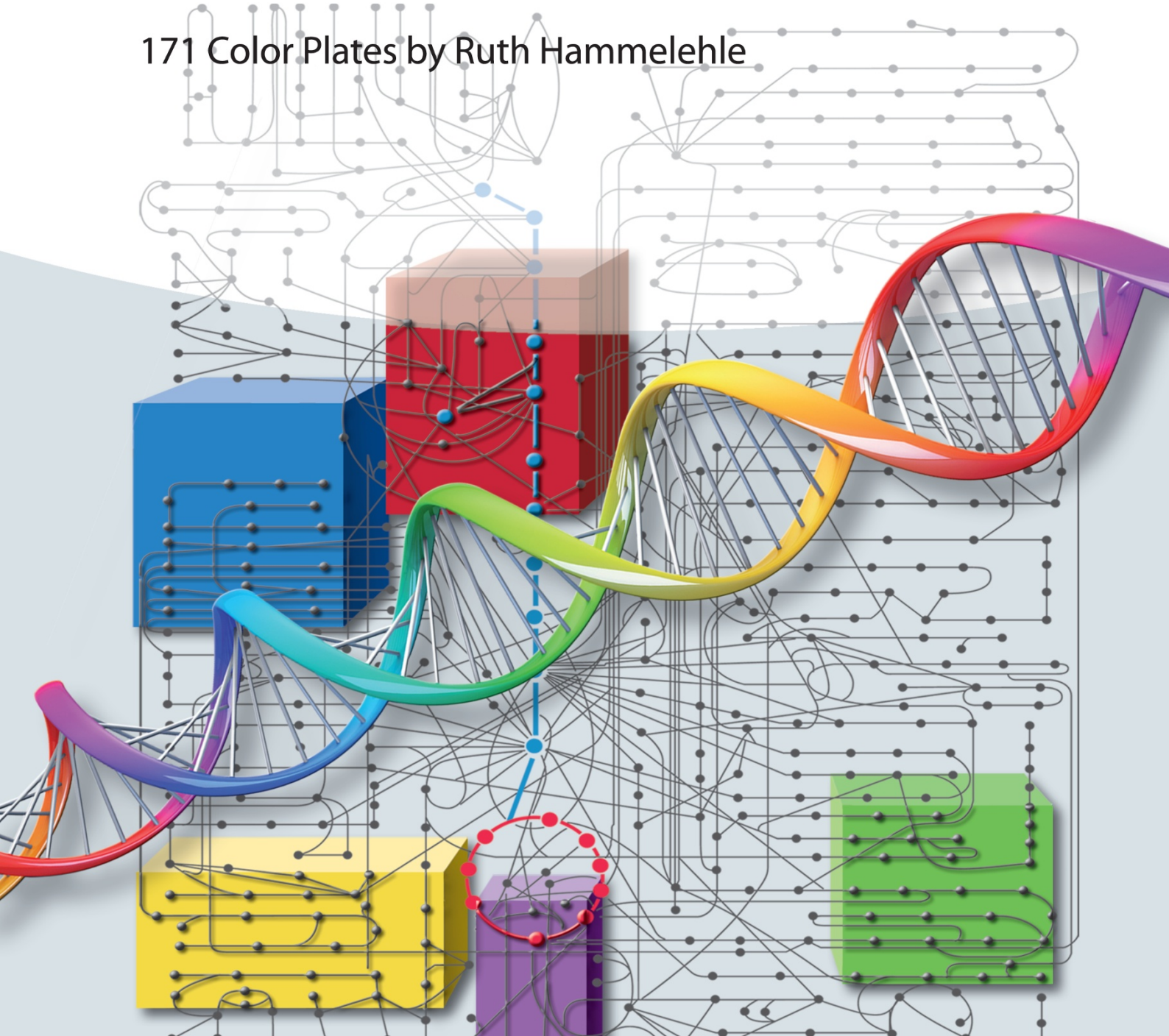
WILEY-VCH

Rolf D. Schmid and Claudia Schmidt-Dannert

Biotechnology

An Illustrated Primer

171 Color Plates by Ruth Hammelehle



Biotechnology
An Illustrated Primer

Rolf D. Schmid and Claudia Schmidt-Dannert

Biotechnology
An Illustrated Primer

171 color plates by Ruth Hammelehle

WILEY-VCH

Authors:

Prof. Dr. Rolf D. Schmid
Bio4Business
Jagdweg 3
70569 Stuttgart
Germany

Prof. Dr. Claudia Schmidt-Dannert
University of Minnesota
Department of Biochemistry
1479 Gortner Ave
140 Gortner Lab
St. Paul, MN 55108
USA

Graphic Designer:
Ruth Hammelehle
Marktplatz 5
73230 Kirchheim unter Teck
Germany

Cover:
DNA helix from fotolia
©A-Mihalis

Library of Congress Card No.:
applied for

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <http://dnb.d-nb.de>.

Print ISBN: 978-3-527-33515-2

Cover Design Adam Design, Weinheim, Germany
Typesetting epline, Kirchheim/Teck, Germany
Printed on acid-free paper

Contents

Megatrends

Preface to the 1st edition

Biotechnology, a key technology of the 21st century, is more than other fields an interdisciplinary endeavor. Depending on the particular objective, it requires knowledge in general biology, molecular genetics, and cell biology; in human genetics and molecular medicine; in virology, microbiology, and biochemistry; in the agricultural and food sciences; in enzyme technology, bioprocess engineering, and systems science. And in addition, biocomputing and bioinformatics play an ever-increasing role. Against this background, it is of little surprise that few concise textbooks try to cover the whole field, and important applied aspects such as animal and plant breeding or analytical biotechnology are often missing even from multivolume monographs.

On the other hand, I have experienced during my own life-long studies, and also when teaching my students, how energizing it is to emerge occasionally from the thousands of details which must be learned, to look at a unifying view.

The Pocket Guide to Biotechnology and Genetic Engineering is an attempt to provide this kind of birds-eye perspective. Admittedly, it is daring to discuss each of this book's topics, ranging from "Beer" to "Tissue Engineering" and "Systems Biology", on a single text page, followed by one page of graphs and tables. After all, monographs, book chapters, reviews, and hundreds of scientific publications are devoted to each single entry covered in this book (many of them are provided in the literature citations). On the other hand, the challenge of surveying each entry in barely more than 4000 characters forces one to concentrate on the essentials and to put them into a wider perspective.

I hope that I have succeeded at least to some extent in this endeavor, and that you will find the clues to return safely from the highly specialized world of science, and its sophisticated terms, to your own evaluation of the opportunities and challenges that modern biotechnology offers to all of us.

This English version is not a simple translation of the original version, which was published in German in December, 2001, but an improved and enlarged second edition: apart from a general update of all data, it contains three new topics (Tissue Engineering, RNA, and Systems Biology).

At this point, my thanks are due to some people who have essentially contributed to this book. Above all, I wish to acknowledge the graphic talent of Ruth Hammelehle, Kirchheim, Germany, who has done a great job in translating scientific language into very clear and beautiful graphs. Marjorie Tiefert, San Ramon, California, has been more than an editor: she has caught and expressed the original spirit of this book. My thanks also to the publisher, in particular to Romy Kirsten. Special thanks are due to the many colleagues in academia and industry who have contributed their time and energy to read through the entries in their areas of expertise and provide me with most useful suggestions and corrections. These were: Max Roehr, University of Vienna; Waander Riethorst, Biochemie GmbH, Kundl; Frank Emde, Heinrich Frings GmbH, Bonn; Peter Duerre, University of Ulm; Edeltraut Mast-Gerlach, Ulf Stahl and Dietrich Knorr, Technical University Berlin; Udo Graefe, Hans-Knoell Institute, Jena; Jochen Berlin, GBF, Braunschweig; Allan Svenson, Novozymes A/S, Copenhagen; Helmut Uhlig, Breisach; Frieder Scheller, University of Potsdam; Bertold Hock, University of Munich-Weihenstephan; Rolf Blaich, Rolf Claus, Helmut Geldermann and Gerd Weber, University of Hohenheim; Hans-Joachim Knackmuss, Dieter Jendrossek, Karl-Heinrich Engesser, Joerg Metzger, Peter Scheurich, Ulrich Eisel, Matthias Reuss, Klaus Mauch, Christoph Syldatk, Michael Thumm, Joseph Altenbuchner, Paul Keller and Ulrich Kull, University of Stuttgart; Thomas von Schell, Stuttgart; Joachim Siedel, Roche AG, Penzberg; Rolf Werner and Kerstin Maier, Boehringer-Ingelheim, Biberach; Frank-Andreas Gunkel, Bayer AG, Wuppertal; Michael Broeker, Chiron Bering GmbH, Marburg; Bernhard Hauer and Uwe Pressler, BASF AG, Ludwigshafen; Frank Zocher, Aventis Pharma, Hoechst; Tilmann Spellig, Schering AG, Bergkamen; Akira Kuninaka, Yamasa Corporation, Chosi; Ian Sutherland, University of Edinburgh; Julia Schueler, Ernst & Young, Frankfurt. Among the many members of my institute in Stuttgart who have patiently helped me with the manuscript I wish

to especially acknowledge Jutta Schmitt, Till Bachmann, Jürgen Pleiss and Daniel Appel.
In spite of all efforts and patient cross-checking, it would be a miracle if no unclearness or errors exist. These are entirely the author's fault. I would be most grateful to all readers who will let me know, via the web address www.itb.uni-stuttgart.de/pocketguide, where this book can be further improved.

Rolf D. Schmid
Stuttgart, New Year 2002/2003

Preface to the 2nd edition

In the 10 years since the first edition of this booklet in English, the developments in biotechnology have further accelerated. This is true for the science, which has generated new methods such as synthetic biology, genome editing or high-throughput sequencing of genomes, generating big data which provide us with ever more detailed perceptions of the living world. New applications in industry have followed suit – in the medical sciences, eminent examples are the therapeutic antibodies, iPS-derived stem-cell technologies or a personalized medicine based on SNP analysis and companion diagnostics; in industrial biotechnology, the emerging concepts of a "bioeconomy" based on renewable resources such as biomass, waste or carbon dioxide provide certainly a megatrend. It goes without saying that a little booklet can only provide short sketches for each of these fields. An updated literature suvey attempts to compensate for this shortcoming.
It is my great joy that Professor Claudia Schmidt-Dannert, University of Minnesota, has accepted to join this and future editions as a co-author. This will help to keep the wide information provided in this book as updated as possible in a global setting.
Our sincere thanks go, beyond the individuals mentioned in the first edition, to numerous friends and colleagues who have helped again with their professional knowledge. Our particular appreciation goes to Wolfgang Wohlleben, Tuebingen University; Karin Benz, NMI Reutlingen; Ulrike Konrad, Protagen; Karl Maurer, ABEnzymes, Darmstadt; Bernhard Hauer, Georg Sprenger and Juergen Pleiss, Stuttgart University; Ulrich Behrendt, Munich; Dirk Weuster-Botz, Munich Technical University; Joern Kalinowsky, Bielefeld University; Vlada Urlacher, Düsseldorf University, and Frieder Scheller, Potsdam University.
The high quality of the artwork is due to Ruth Hammelehle, Kirchheim, of the final editing to Bernhard Walter, both of epline Co., Kirchheim u. T. Our deep thanks to both of them, to the editorial team, Dr Gregor Cicchetti, Dr Andreas Sendtko and Dr Claudia Ley at Wiley-VCH in Weinheim, Germany , and to Dr Sarah Perdue and Dr Bradford Condon at the University of Minnesota, St. Paul. The contribution of Dr Alexandra Prowald, who provided an excellent index to this book, is also highly appreciated.

Rolf D Schmid, Claudia Schmidt-Dannert
Stuttgart, Germany and St. Paul, Minnesota, Summer 2015

Introduction

This pocket guide is written for students of biology, biochemistry and bioprocess engineering who are looking for a short introduction to the many different areas where modern biotechnologies are making an impact. It is also intended as a handy reference for teachers, patent attorneys, managers and investors seeking a quick, yet professional answer surrounding an upcoming topic of industrial biotechnology. To this end, specialized knowledge from a wide range of scientific disciplines has been condensed over a total of 171 color plates and further described on the accompanying text page, as well as complemented by a comprehensive survey of the literature. Cross-references provide additional help in jumping from technical applications of biotechnology, for example, to the fundamental science behind the application.

Completely updated and supplemented by many new topics, this second edition retains the modular format, but the structure of the book has been changed. It now begins, after a brief historical survey, with short introductions to the basic fields of modern biotechnology: **microbiology**, **biochemistry**, **molecular genetics**, **cell biology** and **bioprocess engineering**. It is only in the second part that the focus is on applications, such as **food and food additives**, **industrial products**, **enzyme technology** and, most comprehensively, the many contributions of biotechnology to the medical field, including the manufacture of **antibiotics**, biologicals such as **antibodies**, but also in **medical technology**. This section is rounded off with a description of the applications in **agriculture,** such as animal or plant breeding, and in **environmental protection**. The third section of the book deals with the current megatrends in the applied life sciences. These include **genomics** and such post-genomic trends as **personalized medicine,** with **bioinformatics** seen as an answer to current needs in big data processing, but also **cell technology** and **gene therapy,** as well as technologies devoted to building a new so-called **bioeconomy, i. e.** sustainable in energy and material use. The text ends with five chapters devoted to various aspects of **safety and ethics,** including patent and registration-related topics.

The objective of this book is to provide readers with a compact reference on the wide and expanding field of modern biotechnology. We hope that we succeeded not only in offering an attractive and stimulating read, but also in instigating in the reader the desire to dig deeper into this fascinating area of human endeavor.

Early developments

History. The origins of what we call biotechnology today probably originated with agriculture and can be traced back to early history. Presumably, since the beginning people have gained experience on the loss of food by microbial spoilage; on food conservation by drying, salting, and sugaring; and on the effects of fermented alcoholic beverages. As the first city cultures developed, we find documents and drawings on the preparation of bread, beer, wine, and cheese and on the tanning of hides using principles of biotechnology. In Asia, fermented products such as Sauerkraut (China), Kimchi (Korea) or Gari (Indonesia) have been produced for thousands of years. In Europe, starting in the 6th century, the monasteries with their well organized infrastructure developed protocols for the arts of brewing, wine-making, and baking. We owe our strong, alcohol-rich stout beers to the pious understanding of the monks that "Liquida non fragunt ieiunum" (Liquors do not interfere with the chamfering time). Modern biotechnology, however, is a child of microbiology, which developed significantly in the late 19th century. The First and Second World Wars in the first half of the 20th century next probably provided the strongest challenge to microbiologists, chemists, and engineers to establish modern industrial biotechnology, based on products such as organic solvents and antibiotics. During and after this period, many ground-breaking discoveries and developments were made by biochemists, geneticists, and cell biologists and gave rise to molecular biology. At this point, the stage was set for modern biotechnology, based on genetic and cell engineering, to come into being during the 1970s and '80s. With the advent of information technology, finally, modern biotechnologies gave rise to genomics, proteomics and cellomics, which promise to develop into the key technologies of the 21st century, with a host of applications in medicine, food and agriculture, chemistry and environmental protection.

Early pioneers and products. Biotechnology is an applied science – many of its developments are driven by economic motives. In 1864 Louis Pasteur, a French chemist, used a microscope for the first time to monitor the fermentation of wine vs. lactic acid. Using sterilized media ("pasteurization"), he obtained pure cultures of microorganisms, thus laying the foundation for applied microbiology and expanding this field into the control of pathogenic microorganisms. At the start of the 20th century, it occurred to the German chemist Otto Roehm and to the Japanese scientist Jokichi Takamine that enzymes isolated from animal wastes or from cultures of molds might be useful catalysts in industrial processes. Otto Roehm's idea revolutionized the tanning industry, since tanning up to this time was done using dog excrements. In the field of public health, the introduction of biological sewage treatment around 1900 was a milestone for the prevention of epidemics. During World War I, Carl Neuberg in Germany and Chaim Weizmann, a Russian emigrant to Britain and of Jewish origin, developed large-scale fermentation processes for the preparation of ammunition components (glycerol for nitroglycerol and acetone for Cordite). The Balfour declaration and the ensuing foundation of the state of Israel, whose first president Weizmann became, is thus directly linked to an early success in biotechnology. In the postwar period, 1-Butanol, the second product from Weizmann's Clostridium-based fermentation process, became highly important in the USA as a solvent for car paints. The serendipitous discovery of penicillin by Alexander Fleming (1922), much later turned into a drug by Howard Florey, initiated the large-scale production of penicillin and other antibiotics during World War II. As early as 1950, > 1000 different antibiotics had been isolated and were increasingly used in medicine, in animal feeds, and in plant protection. This was accompanied by a rising tide of antibiotic resistance, triggering research on the mechanisms of microbial defense mechanisms. Since 1950, the analytical use of enzymes, later of antibodies, began another important field of modern biotechnology. The first glucose biosensor was introduced by Leland C. Clark in 1954, initiating a concept for blood glucose monitoring which now commands a market of several billion US-$. In the shadow of the 1960s' oil crises and the emerging awareness of overpopulation, the conversion of biomass to energy such as bio-ethanol and of single-cell protein from petroleum or methanol was developed. Now, in 2014, "biorefineries" are under active development.

Biotechnology in early Egyptian drawings

1 beer brewing
2 bread baking
3 hide tanning

early history	sugar-containing juices are fermented to various alcoholic beverages
	sour milk and sourdough products are prepared by lactic acid and yeast fermentation
	hides are bated to leather using reagents such as animal feces
1650	France: Orléans procedure for the preparation of vinegar from ethanol
~1680	The Netherlands: Anthony van Leuwenhoek observes bacteria through a microscope
1856	France: Louis Pasteur separates brewers yeasts from lactic acid bacteria
~1890	France, Germany: Louis Pasteur, Robert Koch develop the first vaccines
1900	Japan: Jokichi Takamine uses α-amylase for starch degradation
1908	Germany: Otto Roehm uses pancreatic trypsin in detergents and for leather bating
1916	UK: Chaim Weizmann develops a fermentation process for acetone, n-butanol
1920	citric acid is industrially produced by surface fermentation using *Aspergillus niger*
1928/29	UK: Alexander Fleming discovers penicillin
1943	USA: Selman Waksman discovers streptomycin
1949	USA: microbial transformation of steroids on industrial scale
1957	Japan: glutamic acid is industrially produced by tank fermentation of *Corynebacterium glutamicum*
1960	Denmark: *Bacillus* proteases are used in detergents
1965	Denmark: microbial rennet for cheese production
1970	USA: high-fructose syrups produced by enzyme technology replace saccharose in softdrinks
1972/73	USA: Stanley Cohen and Francis Boyer develop a procedure for in-vitro recombination of DNA, using plasmid vectors
1975	UK/Switzerland: César Milstein and Georges Koehler prepare monoclonal antibodies using hybridoma cells
1977	recombinant proteins can be manufactured by fermentation using bacteria
1982	first transgenic plants (herbicide resistance) and animals (knockout)
1985	USA: Kary Mullis discovers the polymerase chain reaction (PCR)
from 1990	USA: the human genome project (HUGO) is initiated
1995	transgenic tomatoes (Flavr Savr) are registered as food in the USA and the UK
1995	gene therapy experiments on humans
1996	the yeast genome is completely sequenced
1998	Dolly the sheep is the first cloned animal, a replicate of its mother
1998	over 2 billion basepairs are stored in DNA sequence databases
1999	the Drosophila genome with 1.6 billion bp is completely sequenced in ~ 4 months
1999	human stem cells can be maintained in culture
1999	the sales of recombinant therapeutic proteins exceed 10 billion US$/yr
2001	Craig Venter's Celera and the international Human Genome Consortium (HGP) present a sketch of the human genome
2008	the USA produces over 30 billion L of bio-ethanol from corn
2012	Shinya Yamanaka, Japan, receives the nobel prize for transforming differentiated cells into autologous stem cells (iPS technology)
2014	transgenic plants are grown on >180 million ha land in 28 nations

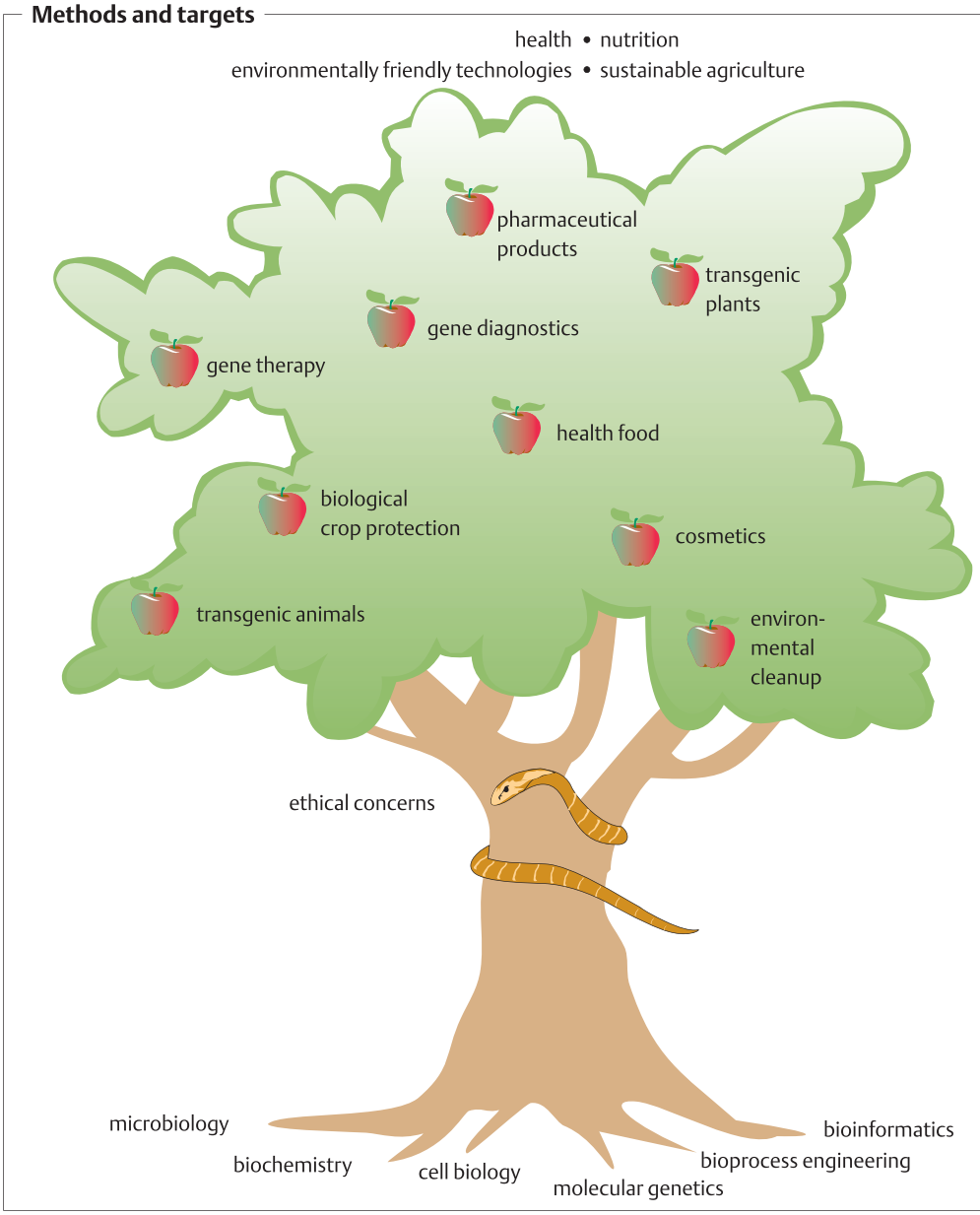
Biotechnology today

Genetic Engineering and Cell Technology. In 1973, Stanley Cohen and Frederick Boyer in San Francisco were the first to express a designed foreign gene in a host organism. After about 10 years the first recombinant drug, human somatotropin, was registered. Since then, more than 50 genetically engineered proteins have been registered as therapeutic agents, including insulin (for diabetics), erythropoietin (for anemic patients), factor VIII (for hemophiliacs), interferon-β (for multiple sclerosis patients), recombinant antibodies and vaccines. Many hundred more are under development. Although the new technologies were first applied to medicine, their innovation potential in agriculture and food production soon began to emerge. Thus, transgenic crops were bred that were resistant to herbicides, insects, or viruses. Today, they are predominantly grown in North America. Flowers have been genetically modified to exhibit new colors, vegetables or fruits to show enhanced nutritive properties, and woods to contain less lignin for improved paper production. In the chemical industry, biopolymers, prepared from biomass-derived chemicals such as starch or cellulose, have begun to replace petrochemical products, and "biorefineries" have appeared which generate biofuels and chemicals from biomass. These technologies are changing the face of agriculture. High-throughput gene sequencers and supercomputers are making the sequencing of human genomes relatively cheap and routine, and genome-based information is now widely used to understand the molecular basis of diseases and to develop novel drugs by a target-oriented screening approach. Novel approaches, such as proteomics and structural biology, are contributing to our fundamental understanding of the chemistry of life and disease. Using gene therapy, we attempt to replace malfunctioning with correctly functioning genes. These developments are in step with great advances in cell biology, which focus on the complex interactions of cells in a multicellular organism. Human differentiated cells such as cardiomyocytes or neurons can now be obtained from embryonal stem cells or even from adult human cells by genetic reprogramming via induced pluripotent stem cells (iPS). Tissue engineering has become a surgical approach to repairing wounded tissue such as skin, bone or cartilage.

Public acceptance. The sheep Dolly, born in 1998, was the first animal ever cloned from a somatic cell of and thus identical to her mother. The thrust and possible consequences of such developments, e. g., for embryonic manipulations or individual (prenatal) genetic fingerprinting, have led to emotional public discussion. Typical subjects are: at what stage does human life begin and when does it need to be protected? Do we accept the cloning of humans? To which extent can we accept a deterministic view of individual health risks, e. g., by an employer or an insurance company? How will molecular genetics and gene therapy affect the age distribution in our societies? Is it ethical to genetically modify plants and animals at will? To what extent are such manipulations in harmony with the ecosystem and its natural diversity? How will the new biotechnologies influence the relationship of industrialized and developing economies? None of these questions has been completely resolved yet. As we begin to understand and interfere with the functions of the human brain, answering these questions on a global scale will become even more urgent.

Foundations. The body of this pocket guide is devoted to the many and growing applications of biotechnology, including discussion of today's "megatrends" (2014), which include bioinformatics. In the introduction to this book, however, the multidisciplinary foundations of the field are briefly outlined. We start with *microbiology*, which is the oldest discipline and has led the way to many contemporary technologies. This is followed by *biochemistry*, the science of life's building blocks, their metabolism and its regulation. A key property of life is to propagate. As a consequence, the basics of *molecular genetics* and *genetic engineering* will be presented. *Cell biology* and *immunology* continue to have a great impact on biotechnology, and some basics are introduced. Finally, without *bioprocess engineering*, a discipline mastered by engineers, the manufacturing of bioproducts could not be done in an economical way. It is obvious that the space available does not allow a thorough discussion of all these fields, but current literature will be provided to the reader interested in further reading.

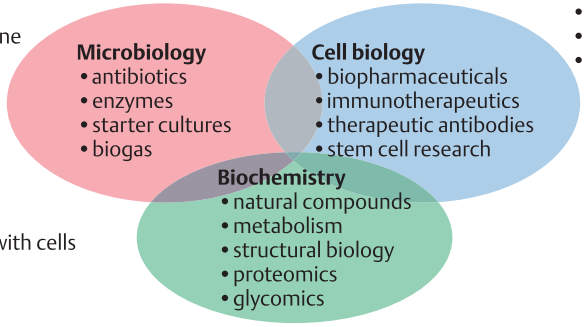
Methods and targets



Scientific foundations

Gene technology/genome research
- recombinant products
- personalized medicine
- gene diagnostics and gene therapy
- animal and plant breeding
- synthetic biology

Bioprocess engineering
- production technology with cells or enzymes
- waste water treatment
- bioenergy



Bioinformatics
- databases
- analysis of "big data"
- metabolic engineering
- systems biology

Viruses

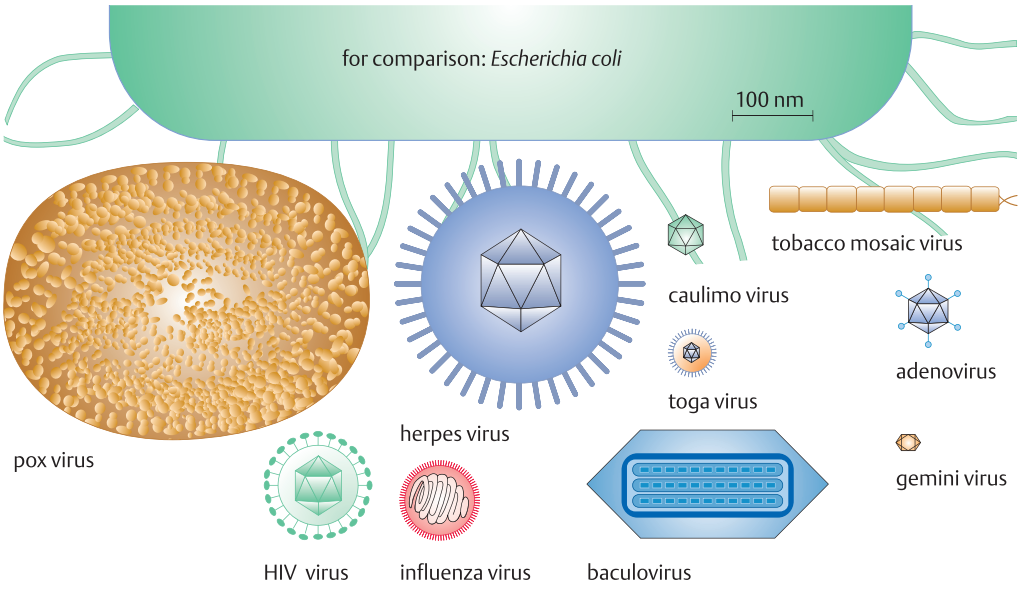
General. A virus is an infectious particle without indigenous metabolism. Its genetic program is written in either DNA or RNA, whose replication depends on the assistance of a living host cell. A virus propagates by causing its host to form a protein coat (capsid), which assembles with the viral nucleic acid (virus particle, nucleocapsid). Viruses can infect most living organisms; they are mostly host-specific or even tissue- or cell-specific. Viruses are classified by their host range, their morphology, their nucleic acid (DNA/RNA), and their capsids. In medicine and veterinary medicine, the early diagnosis, prophylaxis and therapy of viral human and veterinary diseases plays a crucial role. AIDS (HI virus), viral hemorrhagic fever (Ebola virus), avian flu (H5N1-, H7N9-virus) (→250) or hepatitis (several virus families) are important examples of viruses involved in human diseases, as are Rinderpest (Morbillivirus) or infectious salmon anemia (ISA virus) in epizootic veterinary diseases. In biotechnology, viruses are used for the development of coat-specific or component vaccines and for obtaining genetic vector and promoter elements which are, e. g., used in animal cell culture and studied for use in gene therapy.

Viruses for animal experiments. The first cloning experiments with animal cells were done in 1979, using a vector derived from simian virus 40 (SV40) (→98). This virus can infect various mammals, propagating in lytic or lysogenic cycles (lysis vs. retarded lysis of host cells). Its genome of ca. 5.2 kb contains early genes for DNA replication and late genes for capsid synthesis. Expression vectors based on SV40 contain its origin of replication (ori), usually also a promoter, and a transcription termination sequence (polyA) derived from the viral DNA. For the transfection of mouse cells, DNA constructs based on bovine papilloma virus (BPV) are preferred. In infected cells, they change into multicopy plasmids which, during cell division, are passed on to the daughter cells. Attenuated viruses derived from retro, adeno, and herpes viruses are being investigated as gene shuttles for gene therapy (→304). Retroviruses, e. g., the HI virus, contain an RNA genome. They infect only dividing cells and code for a reverse transcriptase which, in the host cell, transcribes the RNA into cDNA. HIV-cDNA is then integrated into the host genome where it directs, via strong promoters, the formation of viral nucleic acid and capsid proteins. Some hundred experiments with retroviral vectors having replication defects have already been carried out for gene therapy. A disadvantage of using retroviral vectors lies in their small capacity to package foreign DNA (inserts), whereas vectors derived from adenoviruses can accommodate up to 28 kb of inserted DNA. In contrast to retroviruses, adenoviruses can infect non-dividing cells, but their DNA does not integrate into the host chromosomal DNA. For gene therapy targeted to neuronal cells, e. g., in experiments related to Alzheimer's or Parkinson's disease, *Herpes simplex*-derived vectors are often used. Their large genome of 152 kb allows them to accommodate inserts > 20 kb of foreign DNA. A similar insert size is reached with Vaccina viruses, which may infect a wide range of cell types.

Viruses for plant experiments. Most plant viruses have an RNA genome (→280). Only two groups of DNA viruses are known that infect higher plants, caulimo virus and gemini virus. Caulimo viruses have a quite narrow host range: they infect only crucifers such as beets and some cabbage varieties. Their small genome reduces their potential for accommodating foreign inserts. Gemini viruses infect important agricultural plants such as maize and wheat and thus bear significant risks for application. Moreover, their genomes undergo various rearrangements and deletions during the infection cycle, rendering the correct expression of foreign DNA inserts difficult.

Baculoviruses infect insects but not mammals. After infection, host cells form a crystalline protein (polyhedrin), which may constitute > 50 % of the insect cell. The polyhedrin promoter is therefore useful for the heterologous expression of proteins, using cell cultures of *Spodoptera* (a butterfly). An advantage of this system is that posttranslational glycosylations in this system resemble those of mammalian cells (→262). Scale-up of this system is, however, limited, rendering it most useful for laboratory experiments. In Japan, silk worms (*Bombyx mori*) are considered an interesting system for expressing foreign proteins. The nuclear polyhedral virus BmNPV is being used for their transfection.

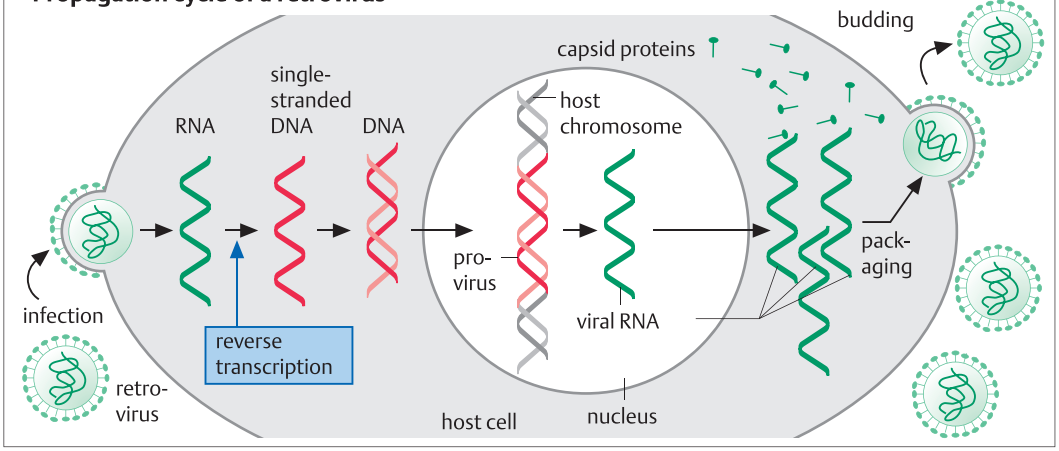
Forms



virus	host	disease	capsid	nucleic acid
smallpox	man, cattle	smallpox	complex coat	linear DNA, d
hepatitis B	man	hepatitis B	polyhedric capsid	circular DNA, d
toga	man	measles	polyhedric capsid	(+)-RNA, s
herpes	man, birds	belt rose, herpes	polyhedric capsid	linear DNA, d
HIV	man, primates	AIDS	round capsid	2 × (+)-RNA, s
influenza	man, mammals	influenza	helical coat	(–)-RNA, segmented
adenovirus	man	common cold	polyhedric capsid	linear DNA, d
papilloma	cattle	warts	polyhedric capsid	circular DNA, d
tobacco mosaic	tobacco plant		polyhedric capsid	RNA, s
caulimo	cabbage		polyhedric capsid	circular DNA, s
gemini	dicots		double polyhedron	circular DNA, s
baculo	insects		polyhedric capsid	circular DNA, d

s = single strand, d = double strand, + = sense direction, – = antisense direction

Propagation cycle of a retrovirus



Bacteriophages

General. Viruses that attack bacteria are termed bacteriophages or simply phages. Their taxonomy is determined by the International Committee on Taxonomy of Viruses, ICTV. Phages occur everywhere in nature, and are abundant in metagenomic analyses of water samples (→74). As there are historic reports of healing by "holy waters," they have been widely studied for the treatment of bacterial infections but results are equivocal. Fermentation processes, e.g., starter culture production (→114), are always endangered by phage infections. As a preventive measure, attempts are usually made to select phage-resistant strains. Phages are useful in genetic engineering, e.g., for the development of cloning vectors or promoters, for DNA sequencing, and for the preparation of gene and protein libraries (→62, 64, 68). Since most cloning experiments use E. coli, phages specific for this bacterium (λ-, M13-, Qβ-, T-phages) play a key role.

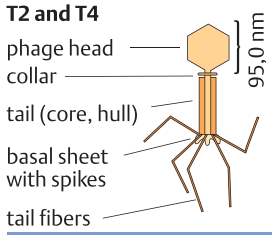
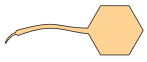
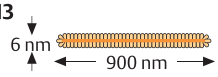
λ Phage. When infecting E. coli, λ phage can follow two routes: either its linear double-stranded DNA (ca. 48.5 kbp, ca. 1% of the E. coli genome) is propagated independent of the E. coli genome, resulting in lysis (lytic cycle), or it is integrated into the E. coli genome, resulting in lysogenic cells containing latent prophages, which replicate with the bacterium over several generations. Upon stress, such as a rise in temperature or UV irradiation, the prophage is excised from the E. coli genome and lyses the host cell. λ is able to form cohesive or sticky ends of 12 unpaired nucleotides each (cos sites), which are necessary for circular λ DNA formation and for its integration into the E. coli genome. The sticky ends also form the recognition signal for the formation of the viral gene product A, an exonuclease. After replication of the λ DNA into a concatemer of linear λ genomes, endonuclease A cuts at this position, initiating the packaging of progeny into its capsids. Cosmids, an important tool for the construction of large gene libraries, are derived from the λ phage, as is a family of λ plasmids such as λEMBL4, which can be induced by a rise in temperature.

The M13 phage infects E. coli according to a different mechanism. It contains single-stranded DNA of ca. 6.4 kb, which after infection directs the synthesis of its complementary strand. The resulting double-stranded phage DNA is not integrated into the E. coli genome but is continuously replicated in the cytoplasm, giving rise to up to 1,000 phage particles/cell. During host cell division, the phage infection is passed on to the daughter cells (ca. 100/cell). Genes that have been cloned into a vector derived from M13 can be obtained as single-stranded DNA – a technique used for classical DNA sequencing (→56). Prior to the invention of PCR, M13 vectors were used for site-directed mutagenesis of proteins.

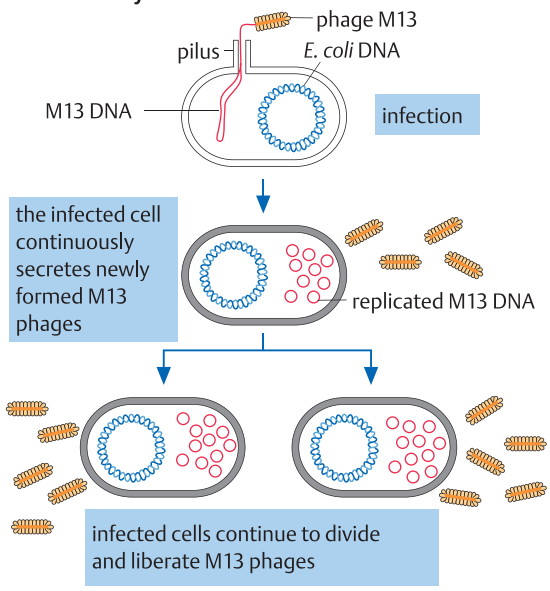
T Phages occur in 7 different types. For genetic engineering, two enzymes coded by T phage genomes are useful: the DNA ligase of T4, which links DNA fragments regardless of the quality of their ends (sticky or blunt), and the DNA polymerase of T7, which polymerizes DNA on a single strand DNA matrix; it is used in gene sequencing (Sanger–Coulson method). The promoter of the T7-RNA polymerase is used in several E. coli expression vectors. T7-RNA polymerase transcribes DNA into RNA, which in turn serves as mRNA in cell-free protein synthesis, based on mRNA, tRNAs, ribosomes, amino acids and ATP.

Phages of other bacteria. Among the > 1,000 classified phages (some 2800 in total), > 300 are specific for enterobacteria, > 230 for bacteriococci, and > 150 each for Bacilli and Actinomycetes. Another group (at present 13 phages), described only recently, is the Ligamenvirales which attack archaeabacteria. Their structure and function are closely related to those of other viruses, including those specific for E. coli. Some of them can be either virulent or lysogenic, similar to the λ phage. Lactobacilli-specific phages are a major problem in the manufacture of milk products. Resistant bacteria prevent adsorption or replication of these phages. Among the 5 groups of Bacillus phages, ø105 and SPO2 are often used for transformation, and PBS1 has been used in construction of the B. subtilis genome sequence map. Phage D3112 is the preferred vector for the transformation of Pseudomonads, and SH3, SH5, SH10, or øC31 are preferred for the genetic engineering of Streptomyces.

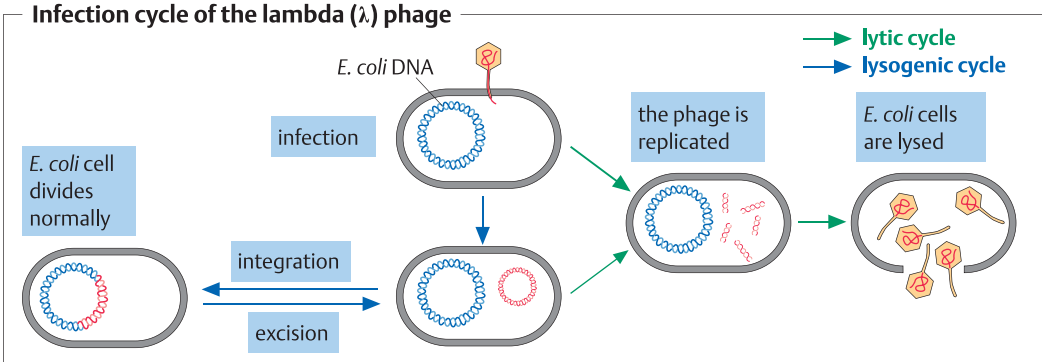
E. coli phages (select)

name	form	genetic material
T2 and T4		DNA (double-stranded)
T7		DNA (double-stranded)
lambda (λ)		DNA (double-stranded)
M13		DNA (single-stranded)

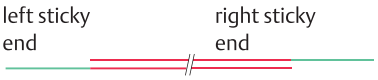
Infection cycle of M13



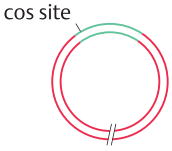
Infection cycle of the lambda (λ) phage

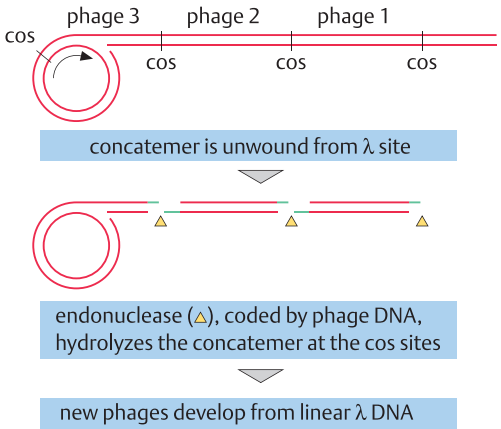


λ DNA in linear form

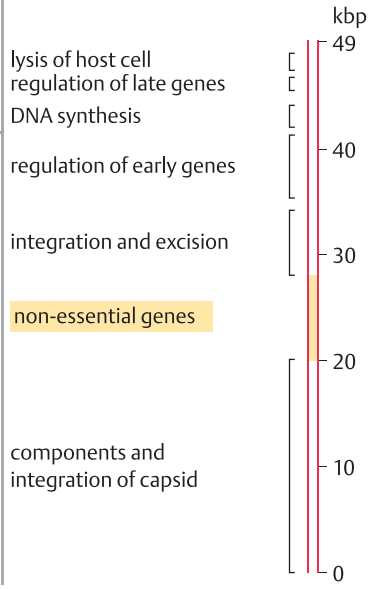


λ DNA in circular form





Genomic map of the λ phage



Microorganisms

General. Microorganisms play a key role in the chemical cycles on earth. They are involved in the biodegradation of many compounds; these processes occur not only in the environment, but also in symbiosis with other organisms (e. g., lichens, intestinal and rumen bacteria). Some microorganisms are parasites or pathogens, impairing the health or life of other organisms. In biotechnology, nonpathogenic microorganisms are used to produce various products such as citric and glutamic acid, antibiotics, xanthan, and enzymes; for the aerobic and anaerobic treatment of wastewater, sludges, soils, and air; and as host organisms for the manufacture of recombinant proteins. Due to their unicellular structure, well established methods for creating and selecting mutants, and their short generation time, they serve as model organisms for understanding the biochemical, genetic, and physiological mechanisms of life, and as a preferred host for the manufacture of recombinant proteins. Based on some fundamental differences, prokaryotic and eukaryotic microorganisms can be distinguished; the former are further subdivided into eubacteria and archaebacteria (> 10,000 different fully characterized strains).

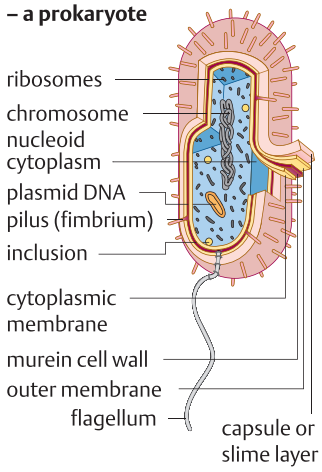
Eubacteria are unicellular organisms that propagate by cell division. Their cell diameter is usually on the order of 1 μm. They have no cell nucleus, and their chromosomal DNA is formed into a tangle, the nucleoid. Frequently, part of their genetic makeup occurs on nonchromosomal genetic elements, the plasmids (→44). Plasmids are often horizontally transferred to other bacteria – a useful mechanism, from the human perspective, for evolving biodegradation pathways for xenobiotic compounds in the environment and sewage plants, but a very dangerous capacity with respect to the evolution of antibiotic resistances. The cell wall, made of peptidoglycan, is more complex in Gram-negative microorganisms and often covered with a slimy layer from which flagella may protrude, which ensure mobility. In the cytoplasm, storage chemicals such as polyhydroxybutyric acid polyphosphate, cyanophycin, or others may be deposited. Eubacteria have a wide potential for variations in metabolism and thus can grow in a much wider range of habitats than higher organisms. Such highly specialized species often surprise us by their unique proteins and cofactors. Thus, the purple membrane of the halobacteria is a unique functional unit of this genus, exhibiting some analogies to photosynthesis and the chemistry of vision in higher organisms.

Archaebacteria (archaea) are believed to resemble the oldest forms of life on earth. Their footprints have been detected in geological formations many hundreds of millions of years old. They often live anaerobically and are usually specialized for growth in unique biotopes. As just one example, the methanobacteria form the most important group of sludge consortia, reducing acetic acid to methane (→288). They differ from the eubacteria in structural and genetic properties, e. g., in the construction of their cell membrane from ether lipids instead of phospholipids. The function of their enzymes is adapted to their often extreme habitats and have been used in biotechnology. For example, a DNA polymerase from a deep-sea bacterium, Pyrococcus furiosus, is often used for PCR reactions with particular high fidelity (→50, 196).

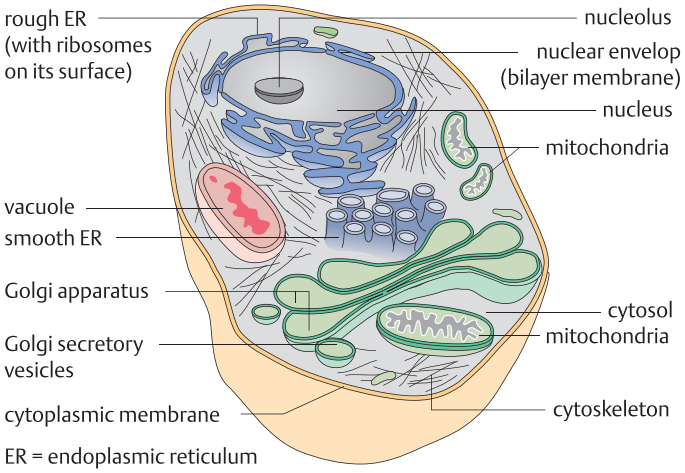
Yeasts and fungi are eukaryotic organisms and so far constitute the largest group of cultivatable microorganisms: about 70,000 different strains have been taxonomically classified. In contrast to prokaryotes, they contain a cell nucleus and other subcellular functional units, and their cell wall is made of chitin, sometimes also from cellulose. Most yeasts and fungi live aerobically. Their wide differences in reproduction and life cycles provide the most useful basis for their taxonomic classification. The vegetative body of fungi is composed of a hairy network, the mycelium, which can propagate sexually or asexually. Asexual reproduction usually proceeds by spore formation, or occasionally by budding. Sexual reproduction of the lower fungi (Phycomycetes) proceeds via gametes, of the higher fungi via fruiting bodies (asci) which have the form of a sac (Ascomycetes) or a club (Basidiomycetes). Yeasts (e. g., Saccharomyces cerevisiae, Schizosaccharomyces pombe) and fungi (e. g., Aspergillus oryzae, Trichoderma viride) are frequently used hosts for the manufacture of recombinant enzymes and other proteins. Unlike prokaryotic hosts, they perform post-translational modifications such as glycosylation (→262), an often important feature for the production of pharmaceutical proteins (glycobiology).

Microorganisms

***Escherichia coli* – a prokaryote**

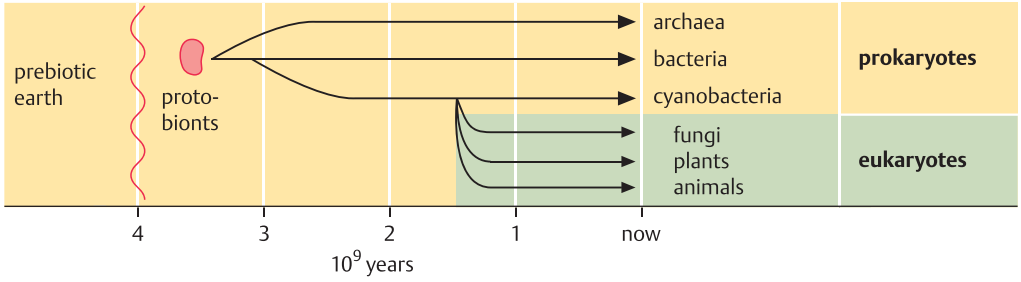


***Saccharomyces cerevisiae* – a eukaryote**



	E. coli	*S. cerevisiae*	for comparison: plant and animal cells
cell nucleus, organelles	no	yes	yes
diameter [µm]	~ 1	~ 10	~ 100
volume [μm^3]	~ 1	~ 1000	>10 000
respiration [µL O_2/mg TS • h]	1000	100	10
generation time [h]	0.3	1.5	> 20
genes	~ 4 300	~ 6 000	> 30 000

Position of the microorganisms in evolution



Archaea, eubacteria, and lower eukaryotes

	archaea	eubacteria	fungi, yeasts
cell type	prokaryote	prokaryote	eukaryote
cell wall	heteropolysaccharide or glycoprotein	peptidoglycan	glucan, chitin
membrane lipids	ether lipids from isoprenoid building blocks	phospholipids	phospholipids
initiator tRNA	methionine	formyl methionine	methionine
genetic material	small circular chromosome, plasmids, histone-type proteins	small circular chromosome, plasmids	complex nucleus with > 1 chromosome and linear DNA, histones
RNA polymerase	complex	simple	complex
size of ribosomes	70S	70S	80S

Bacteria

General. Bacteria can be classified by a variety of morphological, biochemical, and genetic methods, as well as by their nutrient requirements. The *International Code of Nomenclature of Bacteria* (ICNB) governs the scientific naming of bacteria and presently includes about 2,200 genera and 11500 species. The analysis of taxonomically relevant DNA isolated from soil seems to indicate, however, that the number of bacterial species that have not yet been cultured is much larger (→74).

Eubacteria. The oldest method of classifying eubacteria is based on their morphology. Under a simple light microscope, rods, cocci, and spirilli can be seen, some of them forming multicellular aggregates (filaments, colonies) and exhibiting structural details such as spores or flagella. Staining provided further differentiation. Thus, staining according to H. C. Gram's method allows for a classification according to cell wall structure: Gram-positive bacteria have only one cell membrane, covered by a thick murein cell wall, whereas Gram-negative bacteria have two cell membranes, enclosing a periplasmic space. The outer membrane is covered by a thin murein cell wall from which lipopolysaccharides may protrude. Physiological and biochemical criteria have led to additional methods of differentiation. Some important features are:

Response to oxygen: microorganisms can be subdivided according to their ability to grow under aerobic, anaerobic, or both conditions,

Form of energy generation: energy can be generated by photosynthesis (phototrophs), respiration, or fermentation (chemotrophs),

Preferred electron donors: organotrophic microorganisms use organic compounds, and lithotrophic microorganisms use inorganic compounds such a H_2, NH_3, H_2, S, CO, or Fe^{2+}.

Carbon source: autotrophic microorganisms can fix CO_2; heterotrophic microorganisms obtain carbon from organic compounds,

Relation to other organisms: saprophytic microorganisms are autonomous; parasitic microorganisms depend on a host organism.

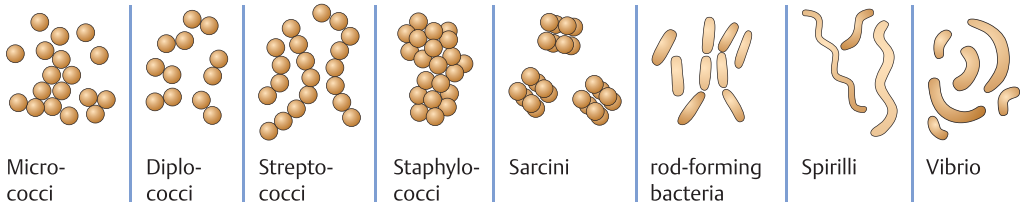
Phage typing: the susceptibility to phages can also be used for taxonomic identification,

Adaptation to environment: mesophilic microorganisms grow under ordinary conditions, whereas extremophiles are adapted to extreme conditions of temperature, pressure, pH, or electrolyte concentration. Cell inclusions, pigments, chemical components of the cell wall and cell membrane (fatty acid composition), immunological differentiation of the cell surface (serology), and susceptibility to antibiotics provide further possibilities for phenotype differentiation. Recently, genotyping of bacteria has become more and more important. For example, the GC content of bacterial DNA enables a rough classification. Complete sequencing of microbial genomes enables the most precise differentiation. A particularly useful method for taxonomy, discovered in 1972, is sequencing the DNA coding for the 16S, 18S and 23S rRNA (S: Svedberg units characterizing sedimentation behavior). This DNA contains sequences that were highly conserved throughout evolution, and analyses of the sequences suggest three families of living organisms: archaebacteria, eubacteria (the prokaryotes), and the eukaryotes. If DNA is isolated from environmental samples, and sequences coding for 16S, 18S or 23S rRNA are compared to those of microorganisms deposited in culture collections, there is less than 5 % identity, suggesting that > 95 % of all microorganisms contained in these samples have not yet been cultivated (s. metagenome) (→74).

Characterization and taxonomy. Rapid taxonomic identification of bacteria is important in hospitals, veterinary medicine, food production, environmental hygiene, and also in microbial and genetics laboratories. Most of the above methods are used, e. g., microscopy, staining procedures, determining the "analytical profile index API" (based on growth on various substrates), fatty acid composition of the membrane, or DNA analysis of taxon-specific sequences coding for the 16S, 18S or 23S rRNA. Precise classification of microorganisms is often far from trivial and requires the consideration of a wide range of experimental data; it is usually done by laboratories that archive culture collections.

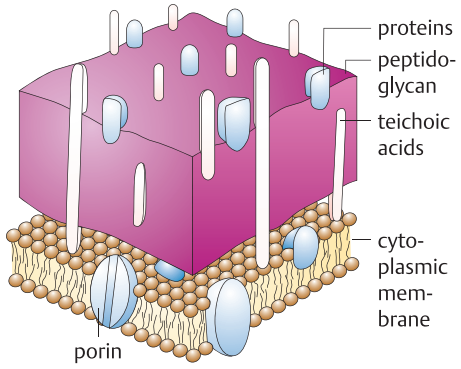
Genome sequencing. As of 2013, genome sequences for some 2,100 bacteria and over 140 archaea are completed. This includes many genomes of human pathogens such as *Mycobacterium tuberculosis*. The analysis of microbial genomes has shown that many variations of metabolic pathways exist, which can be exploited by metabolic engineering.

Forms of unicellular bacteria

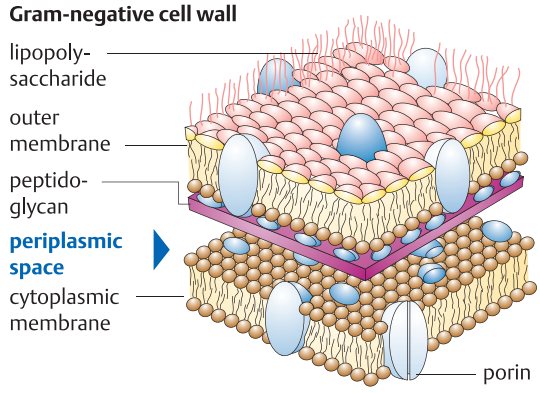


Cell wall composition and Gram-staining

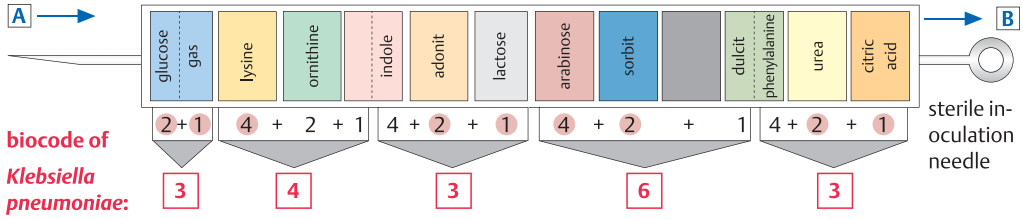
Gram-positive cell wall



Gram-negative cell wall

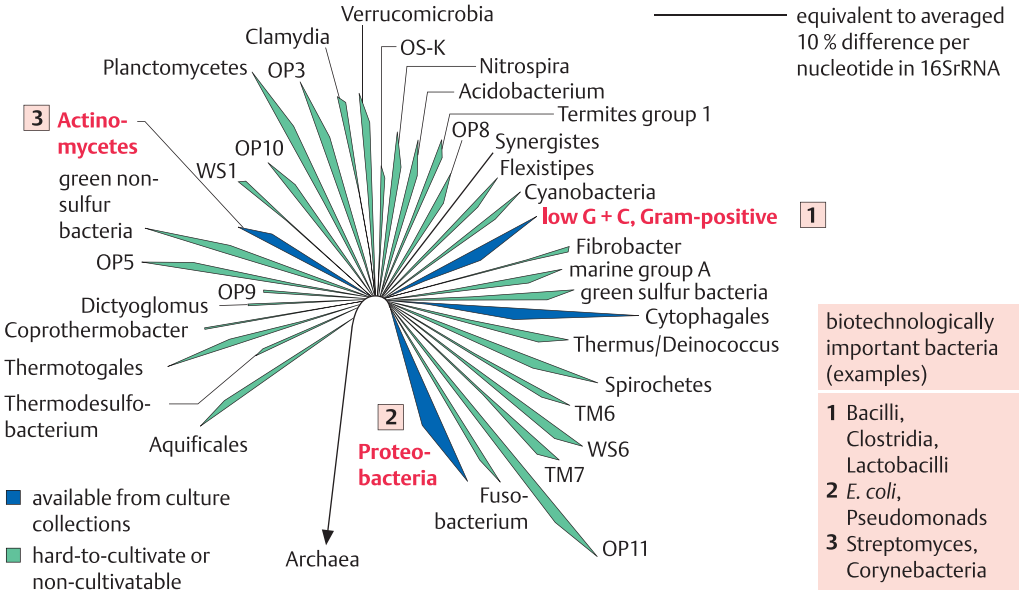


Biochemical characterization



color reactions after inserting the inoculation needle into a clinical probe A, draw through all compartments B, incubate

Phylogenetics and cultivation



Yeasts

General. Yeasts are a subgroup of the Ascomycetes. Because they propagate by budding, they are also termed budding fungi. They grow heterotrophically, preferring acidic media (pH 3.5–5.0) and usually do not form mycelia. Their cell wall is made of chitin. *Candida albicans* is an important human pathogen and model for studying pathogenesis. Yeasts of importance for biotechnology are *Saccharomyces cerevisiae*, *Candida utilis* and other *Candida* strains, *Schizosaccharomyces pombe*, *Hansenula polymorpha*, and *Pichia pastoris*.

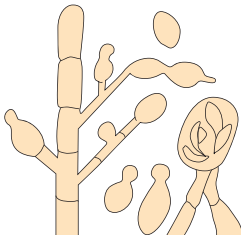
Saccharomyces cerevisiae (synonyms: baker's yeast, brewer's yeast, yeast) (→120) can propagate in either a haploid or diploid manner, thus providing an excellent organism for genetic investigations. Haploid laboratory strains belong to one of two mating types (*MATa* or *MATα*), which can only mate reciprocally. Asexual reproduction proceeds by forming conidia, followed by immigration of either a diploid or a haploid nucleus. Sexual propagation occurs by the fusion of two haploid gametes, followed by meiosis and formation of 4 haploid ascospores, whose phenotype can be separately observed, allowing for simple genetic analysis of the observed traits (tetrad analysis). Due to the simple cultivation of both haploid and diploid cells, the completed genome sequence (12 Mbp, on 16 chromosomes), the general absence of introns, and the short doubling time (90 min), *S. cerevisiae* has become a preferred model organism for the molecular genetics of a simple eukaryote. Another advantage is that yeast occurs with a natural plasmid, termed 2μm (60–100 copies in the cell nucleus), and that a second extrachromosomal element, the killer virion, is also available for recombination experiments. Many cloning vectors have been developed for yeast transformation, which either allow the replication of foreign genes outside the yeast chromosome (YRP = yeast replicating plasmids or YEP = yeast episomal plasmids) or integration of the foreign gene into the chromosome (YIP = yeast integrating plasmids). Artificial yeast chromosomes (YAC = yeast artificial chromosomes) allow for the cloning of large DNA fragments of 600–1,400 kbp; they have been widely used for preparing genome libraries, but have a tendency to recombine and thus have been mostly replaced by bacterial artificial chromosomes (BACs) (→72). The ca. 6,000 genes of yeast, located on 16 linear chromosomes, often show high homology to human genes. Thus yeasts have widely served as a simple model system for metabolic and regulation studies. In biotechnology, yeasts are used in the preparation of food products such as beer (→112), wine (→110), and bread (→120). It is also used in the manufacture of industrial ethanol (→138). Recombinant yeasts have become important host organisms for the manufacture of products such as insulin (→222), interferons (→234), and vaccines (→250) (e. g., hepatitis B surface antigen). Unlike *E. coli*, yeast allows for the posttranslational modification of gene products, in particular for glycosylation (→262).

Candida utilis differs from *Saccharomyces* by forming a mycelium, but it propagates solely asexually by budding. Some *Candida* genes show noncanonical codon usage (e. g., CUG for serine instead of leucine), which has retarded their heterologous expression. *Candida* strains have been used in biotechnology for production of extracellular enzymes and generation of digestible biomass. They can be grown on unconventional substrates such as sulfite suds or alkane fractions. Some *Candida* strains, such as *Candida albicans*, are pathogenic to humans.

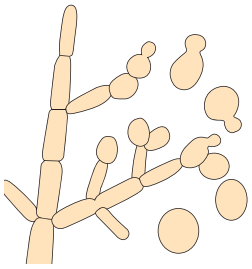
Pichia pastoris and Hansenula polymorpha are methylotrophic yeasts, which can grow on methanol as their sole carbon source. Isolated and studied in the context of the manufacture of single-cell protein (→122), they are used today as attractive host organisms in cloning experiments. Thus, diverse proteins such as lipases, β interferon, and antibody fragments have been functionally expressed in *P. pastories* in yields of several grams of recombinant products/L of culture broth. The *Hansenula polymorpha* genome (9.5 Mbp, 6 chromosomes) was sequenced in 2003, the *Pichia pastoris* genome (9.4 Mbp, 4 chromosomes) was sequenced in 2009.

Schizosaccharomyces pombe was first isolated from an East African beer variety (Swahili: pombe = beer). The genome of this ascomycete was fully sequenced in 2002 (12.6 Mbp, 3 chromosomes), and is similar in size to the *S. cerevisiae* genome. Mutant strains with reduced genome size and partial deletion of protease genes have been constructed which allow for excellent expression of foreign proteins.

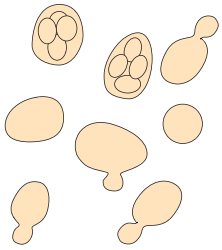
Morphology



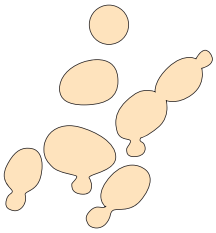
Endomycopsis
forms ascus
forms mycelium
forms buds



Candida
–
forms mycelium
forms buds



Saccharomyces
forms ascus
–
forms buds



Torulopsis
–
–
forms buds

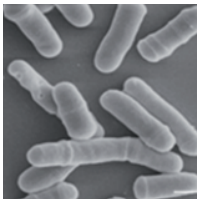
genetic

	size of haploid genome [Mbp]	chromo-somes	gene	genome-sequence
Saccharomyces cerevisiae	12.1	16	5905	1996
Candida utilis	14.6	14	8646	2012
Pichia pastoris	9.4	4	5040	2009
Hansenula polymorpha	9.5	6	5933	2003
Schizosaccharomyces pombe	14.1	3	4970	2002
for comparison: *Escherichia coli* K12	4.6	1	4145	1997

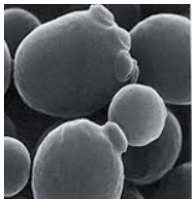
Technical applications of yeasts

Saccharomyces cerevisiae	• baker's yeast, brewers yeast • host organism for the expression of peptides, proteins and enzymes • model organism for the analysis of metabolic and gene regulation • modell organism for aging research
Candida strains	• animal feed • manufacture of biosurfactants • biotransformation reactions
Pichia pastoris, Hansenula polymorpha	• host organisms for the expression of proteins and enzymes
Schizosaccharomyces pombe	• host organism for the expression of proteins and enzymes • model organism for the analysis of gene regulation

Yeasts

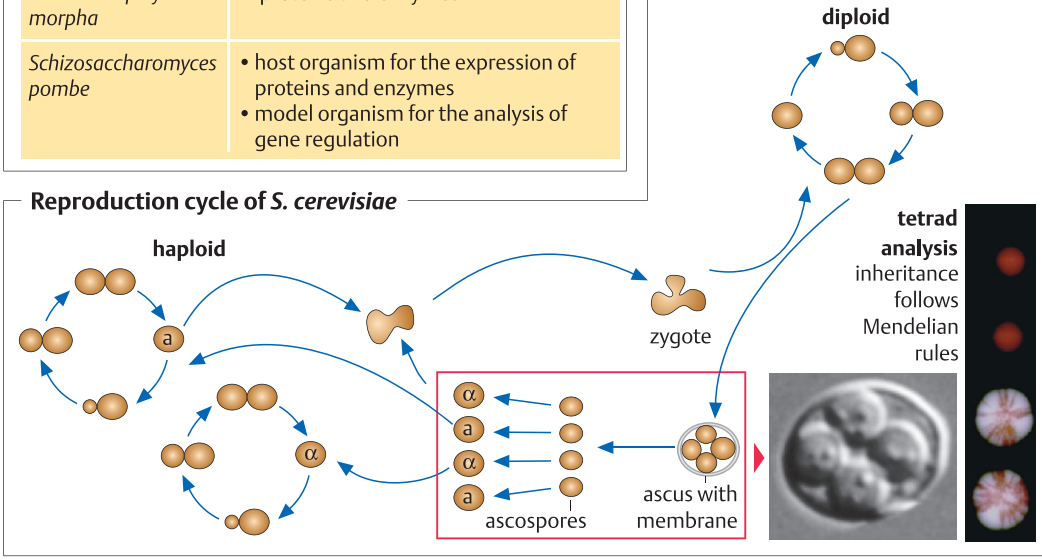


Schizosaccharomyces pombe



Saccharomyces erevisiae
= baker's yeast

Reproduction cycle of *S. cerevisiae*



Fungi

General. Fungi play a key role in the carbon catabolism of the biosphere, e. g., in the decomposition of wood and the formation of humic acids. Mycorrhizal fungi are associated with plant roots and assist in the uptake of nutrients, but other fungi, such as mildews, are dangerous plant pathogens. In biotechnology, they have an important role in the decay of food, but also in the preparation of fermented food products. Some fungi produce antibiotics or valuable enzymes. Among ca. 70,000 fungal species that have been classified, the Ascomycetes comprise ca. 20,000 species, forming the largest subgroup, which includes *Penicillium notatum* and *Aspergillus niger*. Among the lower fungi (Zygomycetes), *Rhizopus* and *Mucor* species have the greatest importance in biotechnology. Some of the ca. 12,000 stand mushrooms (Basidiomycetes) are edible (e. g., champignons, shiitake, chanterelles, ceps), and others participate in the degradation of wood (white and red rot fungi). Approximately 300 fungal species are pathogenic to humans. All fungi live heterotrophically. Their cell wall is composed of chitin and glucans.

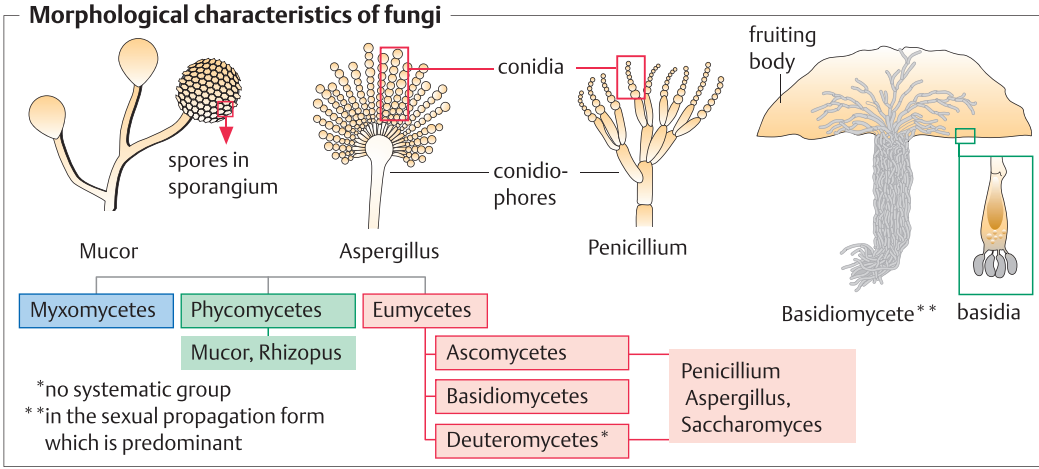
Reproduction forms. The reproduction of fungi follows highly diverse patterns, which are described here using the Ascomycetes as an example. The cell mass (thallus) consists of a mycelium that is made up of hyphae. During asexual reproduction, the conidiophores, which form at the top of the mycelium, divide and form spores (conidia), which grow into a new mycelium. Like most fungi, Ascomycetes can also propagate by a sexual mechanism. This results in a different phenotype (dimorphism). In this case, their hyphae form male and female sexual organs (antheridia and ascogonia). They fuse, during plasmogamy, into dikaryotic hyphae, which develop into an ascocarp ("fruiting body"). In the terminal cells of the dikaryotic hyphae, the dikaryotic nuclei are fused into a diploid zygote (karyogamy). Meiosis transforms the zygote into 8 haploid ascospores (or 4 basidiospores, in Basidiomycetes), which again grow into a mycelium.

Penicillium chrysogenum grows as a mycelium which forms fruiting bodies liberating spores (conidia) for asexual reproduction. Fungi like *Penicillium*, which have lost the capacity for sexual reproduction, are termed *Fungi imperfecti*. Consequently, if recombination is required during breeding in the laboratory, protoplast fusion among different types of nuclei (heterokaryosis) must be used. *P. chrysogenum* and the related fungus *Acremonium chrysogenium* are important industrial organisms, since they synthesize the lactam antibiotics (→206). Other *Penicillium* species such as *Penicillium camembertii* play an important role in the maturation of cheese (→188). The genome of *P. chrysogenum* contains ca. 32 Mbp and the sequence was published in 2008.

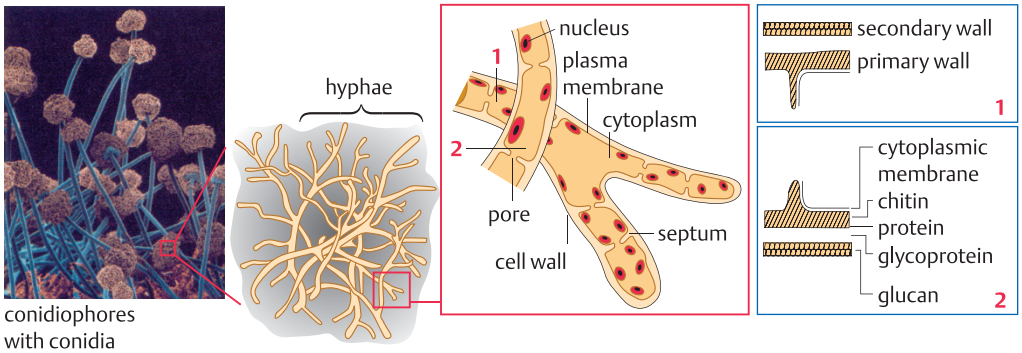
Aspergillus nidulans differs from *Penicillium* in the form of its conidia. Its genome contains 30.5 Mbp. *A. oryzae* is used for industrial production of extracellular enzymes (→172) and is a favorite host organism for producing recombinant enzymes from other eukaryotes. Various *Aspergillus* strains play a traditional role in Asian countries for the manufacture of food products such as soy sauce, miso, and sake (→86, 114), and their genetic and biochemical properties related to the production of these products have been analyzed in great detail. *Aspergilli* are also used for the production of extracellular enzymes such as proteases or amylases, and are preferred hosts for the production of recombinant fungal enzymes which they secrete. *A. niger* is the preferred production organism for citric and gluconic acid (→146, 150). Similar to *Penicillium*, strain improvement still uses protoplast fusion and selection; as the genome sequences of *A. nidulans, A. niger, A. oryzae* and eight more *Aspergillus* strains are now available (2013), targeted strain improvements based on the molecular genetic analysis of desired traits are rapidly advancing.

Rhizopus oryzae, a zygomycete, grows on rice, and *R. nigricans* is the black mold on bread. Its hyphae grow rapidly and bore their way through their substrates. Asexual reproduction proceeds by the formation of spores in differentiated mycelium (sporangia). *Rhizopus* and the closely related *Mucor* species can also grow on decaying organic materials and synthesize numerous extracellular hydrolases for this purpose. As a result, they have become important organisms for the manufacture of extracellular enzymes such as lipases and proteases. The *R. oryzae* genome is composed of 45.2 Mbp and was completely sequenced in 2009. A second *Mucor* genome sequence is available from *Mucor circinelloides*.

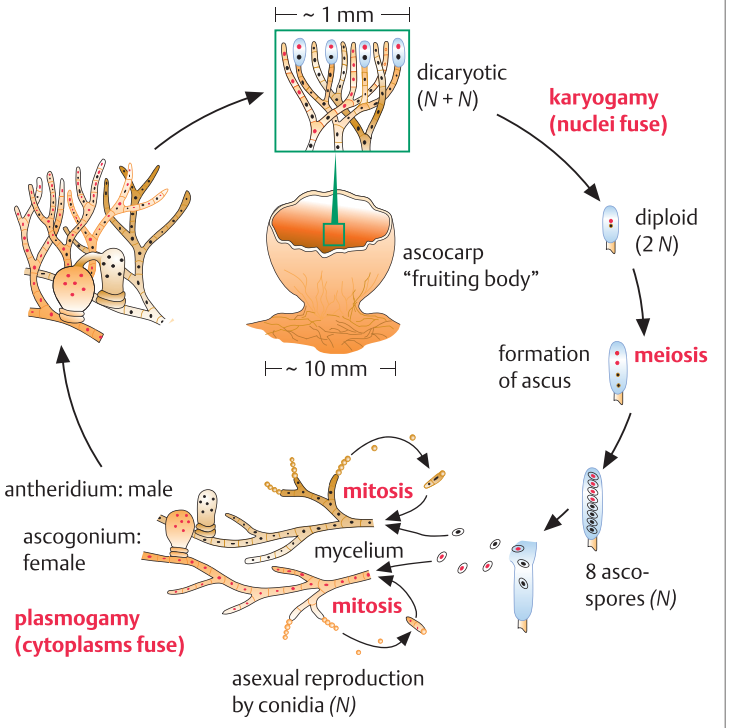
Morphological characteristics of fungi



Aspergillus niger, an ascomycete

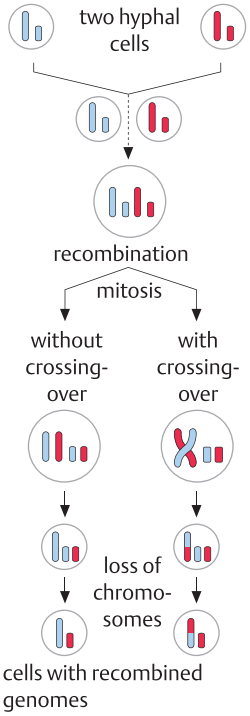


Propagation cycle of an ascomycete



Parasexual breeding

e.g., of Aspergillus



Algae

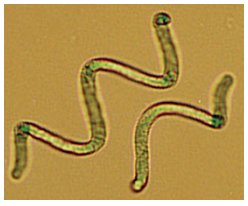
General. Most algae live in water, assimilate CO_2 and produce O_2 through photosynthesis. Unlike terrestrial plants, they do not pass an embryonic stage. Prokaryotic algae are termed cyanobacteria or "blue-green algae" and classified into about 100 genera. Eukaryotic algae form > 20,000 genera and are subdivided into green, brown, red algae, diatoms and others. Some cyanobacteria and algae form toxins (microcystins, saxitoxin), others are used for the production of food additives or specialty chemicals such as, e. g., alginate, agar-agar or astaxanthin. More recently, algae have been explored as a source of bio-energy, since, like terrestrial plants, they use CO_2 as their sole carbon-source, but without competing for agricultural land. Bio-energy may be harvested as algal biomass (formation of biogas) or lipids (triglycerides or isoprenoids), sometimes optimized in yield by metabolic engineering. Cultivation can be performed by aquaculture, in open ponds or in bioreactors. The biotechnology of algae is mainly promoted in nations with plentiful sunshine and long coasts such as the USA, Australia, Japan, Israel and China.

Eukaryotic algae comprise unicellular organisms of some µm size (Chlorella), but also large multicellular organisms up to 30 m long (kelp). Algal cells contain compartments such as chloroplasts which, in addition to chlorophyll a und b, often contain carotenoids. Some genera such as, e. g., ***Euglena***, can live both as autotrophs or complete heterotrophs. In the latter case, they lose their chloroplasts. The cell wall of many algae is composed of cellulose fibrils which are reinforced by other polysaccharides such as alginic acid. Diatoms form their cell wall from silicates built by silica deposition on a protein matrix. ***Laminaria*** and other marine brown algae are an important source of alginates (→158). The viscosity of an alginate solution depends on Ca^{2+} concentration. Alginates are used in the food industry as thickening agents, in medicine for surface wound treatments, and recently also as fibers in textile production. ***Chlorella*** are unicellular freshwater algae which propagate asexually. They contain one chloroplast and only a few mitochondria. Their cultivation is quite simple, and they are used as food additives. ***Botryococcus braunii*** is another green freshwater microalgae. Unlike *Chlorella*, it forms colonies. Under appropriate conditions, it can accumulate up to 60 % hydrocarbon content (alkanes, terpenoids, squalen). *Botyrococcus* oil is being investigated as a biofuel. ***Haematococcus pluvialis*** is a freshwater algae which forms cocci. It is able to synthesize the red-colored tetraterpene astaxanthine in high yields. Through the aquatic food chain, astaxanthine is responsible for the reddish color of salmon, shrimps etc.. It is a strong antioxidant which is well tolerated in human nutrition. As a consequence, it is used as a food additive and in cosmetics. ***Crypthecodinium cohnii*** is a marine red algae from the family dynoflagellatae. It can form up to 20 % of its dry mass as dodosahexaeoic acid (DHA), an ***ω-3***-fatty acid (→34, 162), which is used as an antioxidant food additive. ***Dunalilla*** are halophilic marine microalgae. They form high concentrations of ***ß***-carotene and glycerol, the latter as an osmoregulant. ***Neochloris oleoabundans*** is a green micro-algae which accumulates up to 30 % of its dry mass as triglyceride. This oil is being investigated as a biofuel. The genus ***Nannochloropsis*** comprises several marine phytoplankton algae, some of which store triglycerides. As they are quite easy to transform, they may have potential for producing alkanes from fatty acids, using synthetic metabolic pathways.

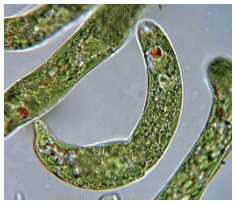
Cyanobacteria are prokaryotic organisms. Some of them can also grow in a heterotrophic manner. They exhibit wide morphological diversity and are divided into 5 classes. Their cell walls are composed of peptidoglycan, and their photosynthetic membranes are multi-layered and complex: besides chlorophyll a, they contain phycobiline pigments. Many cyanobacteria contain "heterocysts" for nitrogen fixation and cyanophycine, an aspartate-arginine copolymer, as a carbon-nitrogen storage compound. The genomes of about 35 cyanobacteria have been sequenced, and the molecular biology of ***Synechocystis sp.*** is most advanced. ***Spirulina*** is a 1–3 µm long cyanobacterium which grows in highly alkaline salt lakes. It forms multicellular spiral microfilaments. Spirulina biomass is produced in aquaculture and marketed as a food and feed additive.

Algae

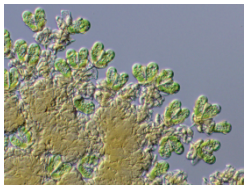
phylogeny	properties	examples relevant for biotechnology
cyano-bacteria	prokaryotes only chlorophyll a cell wall: peptidoglycans	*Spirulina, Synechocystis sp.*
algae	eukaryotes chlorophyll a and b cell wall: cellulose, polysaccharides, silicate	*Euglena gracilis, Chlorella vulgaris, Botryrococcus braunii, Haematococcus pluvialis, Dunaliella salina, Nannochloropsis oculata, Neochloris oleoabundans, Laminaria*



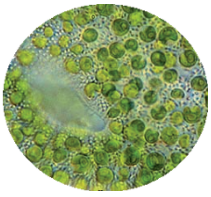
Spirulina



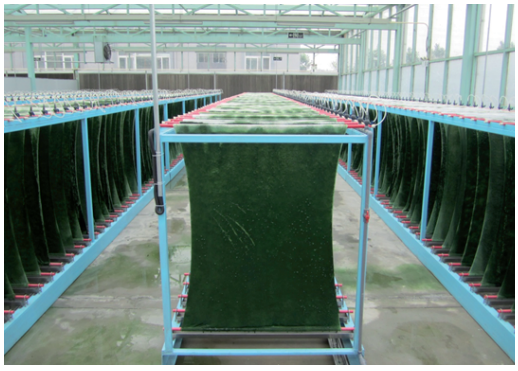
Euglena



Botryococcus braunii



Chlorella

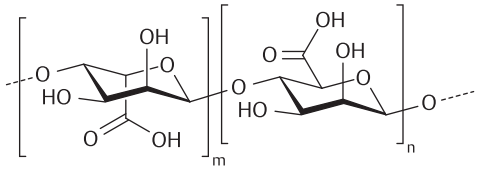


Spirulina platensis cultivation in attached biofilm at CAS Institute for BioEnergy and Bioprocess Technology (QIBEBT), Qingdao, PR China

Industrial manufacture of algal products

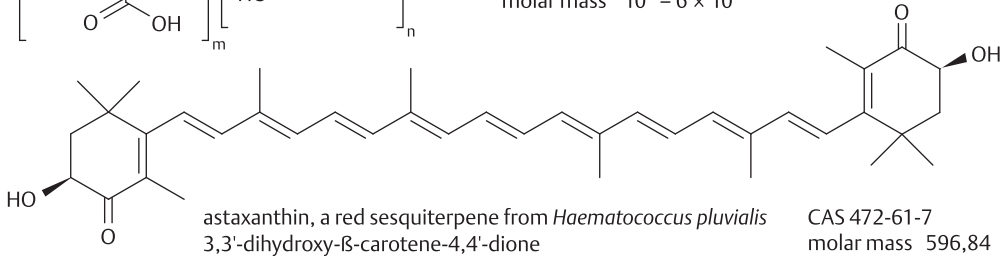
Dunalliela salina Farm of Cognis Co., Western Australia

Neochloris oleoabundans biofuel algae farm, conceptical drawing



Alginate, a copolymer from Laminaria algae
α-L-gulonic acid and β-D-mannuronic acid

CAS 9005_32-7
molar mass $10^4 - 6 \times 10^5$



astaxanthin, a red sesquiterpene from *Haematococcus pluvialis*
3,3'-dihydroxy-ß-carotene-4,4'-dione

CAS 472-61-7
molar mass 596,84

Some bacteria of importance for biotechnology

General. Some bacteria are especially important in biotechnology. Examples are *Escherichia coli*, *Pseudomonas putida*, *Bacillus subtilis*, *Streptomyces coelicolor*, and *Corynebacterium glutamicum*.

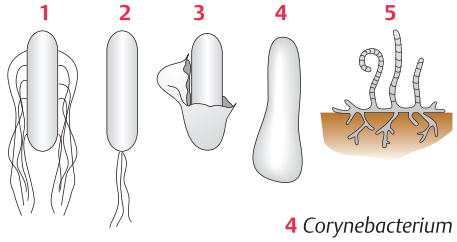
Escherichia coli is a saprophyte in the large intestine of mammals and belongs to the Enterobacteriaceae group. It forms rods that carry flagella. The cell wall stains Gram-negative: it encloses two membranes that include a periplasmic space. Under anaerobic growth conditions, *E. coli* generates energy by fermentation and forms acids. In the presence of O_2, energy is supplied through the respiratory chain. Under optimal conditions, its doubling time is ca. 20 min. The *E. coli* genome is ca. 4.6 Mbp in size, the G+C content is 51 %. Although *E. coli* is among the best understood microorganisms and the genome of *E. coli* K-12 MG1655 was sequenced in 1997, the function of many of its gene products derived from ~4,300 open reading frames (ORFs) is not yet fully understood. In biotechnology, *E. coli* is used as a host organism for the expression of nonglycosylated proteins, e. g., enzymes, insulin, growth hormone, and antibody fragments. Since *E. coli* grows in the human large intestine, it is classified in safety group S2; as a consequence, attenuated *E. coli* strains of reduced genome size are used, in which all risk factors were eliminated and which can be handled under normal microbiological safety conditions as group S1 organisms (e. g., *E. coli* K12) (→332). They are also used for cloning experiments. Various plasmid vectors (→58) have been developed for cloning foreign genes in *E. coli*, for example, the BAC cloning vector is used to construct genomic libraries.

Pseudomonas putida is rod-shaped with polar flagella and lives aerobically in water. The cell wall contains two membranes that enclose a periplasmic space and stains Gram-negative. The *P. putida* genome contains ca. 6.1 Mbp, its G+C content is 61 %. Pseudomonads have a wide genetic potential for the degradation of aromatic compounds, which can be horizontally transferred through plasmids. In biotechnology, they are mostly used in environmental studies (→292). ***Bacillus subtilis*** is rod-shaped without flagella and lives aerobically in soil. Under unfavorable conditions, it forms dormant, thermoresistant spores. Its cell wall stains Gram-positive and encloses only one membrane. Energy is generated via the electron transport chain. Doubling time, under optimal growth conditions, is ca. 20 min. The genome of *B. subtilis* contains ca. 4.2 Mbp and has been completely sequenced; its G+C-content is 44 %. In biotechnology, *B. subtilis* is the preferred microorganism for producing extracellular enzymes, e. g., proteases, cellulases and amylases (→174, 176, 190, 194). It is also used for the production of some antibiotics such as bacitracin. Production strains of 20 % reduced genome size have been engineered which produce up to two-fold more cellulase or protease.

Corynebacterium glutamicum is a member of the coryneform bacteria which grow in many habitats and include some pathogenic species such as *C. diphteriae*. The club-shaped cells grow aerobically and stain Gram-positive. The *C. glutamicum* genome contains ca. 3.1 Mbp and was completely sequenced in 2003; its GC content is 56 %. Deregulated and metabolically engineered mutants of *C. glutamicum* are important production strains for L-glutamic acid and L-lysine. *C. glutamicum* is a preferred organism for synthetic biology (→320), and mutant strains which overproduce lactic acid (→148), succinic acid (→152), 1,2-propanediol (→142) or aniline from biomass have already been described. A Corynex® system based on *C. glutamicum* mutant strains has been proposed for the industrial manufacturing of pharmaceutical proteins in high yields, with excellent down-stream processing.

Streptomyces coelicolor is another soil bacterium from the genus Actinomycetes. It propagates in the form of a mycelium and forms aerial hyphae, from which spore-forming conidia are constructed. The cell wall stains Gram-positive and encloses just one membrane. Like most other Streptomyces strains, *S. coelicolor* degrades cellulose and chitin. Its large linear genome has been completely sequenced and contains ca. 8.7 Mbp, nearly twice the number in *E. coli*; its G+C content is 72 %. The ca. 8,000 structural genes code mainly for enzymes that are required for the formation of secondary metabolites, e. g., for antibiotics (→200).

Some important bacteria in biotechnology

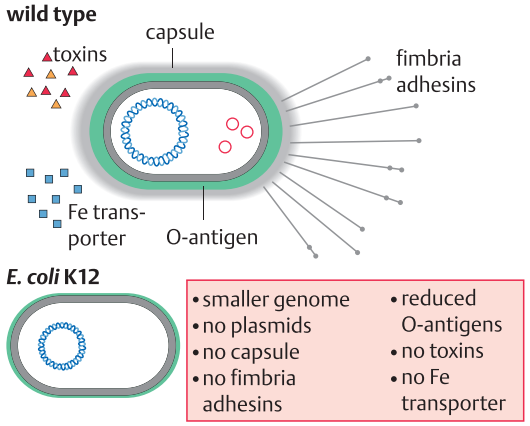


1 *Escherichia coli*
2 *Pseudomonas putida*
3 *Bacillus subtilis* (germinating from spore)
4 *Corynebacterium glutamicum*
5 *Streptomyces coelicolor* (with conidia)

	1	2	3	4	5
flagellation	+	+	–	–	–
Gram-staining	–	–	+	+	+
spore formation	–	–	+	–	+
aerobic growth	+	+	+	+	+
G + C content	51	61	44	56	72
genome size (Mbp)	4,6*	4,2	4,2*	3,1*	8,7*

*genome sequences have been completed

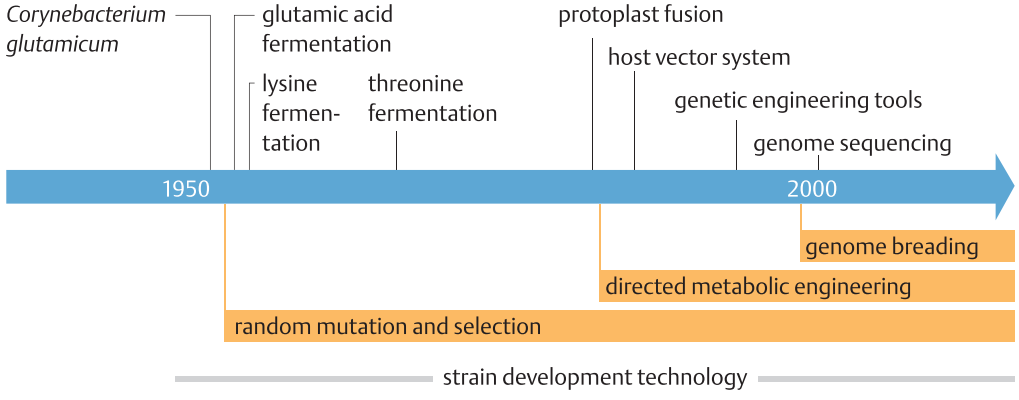
E. coli K12 modifications



E. coli K12 genome: gene functions

total	4485
enzymes	~1500
transport proteins	~600
regulatory proteins	~400
genes of foreign origin	~300
membrane proteins	~250
structural proteins	~200
carrier proteins	~100
RNA synthesis	~150
other	~300
unknown function	~600

Corynebacterium glutamicum



Some completely sequenced genomes of prokaryotes

	disease	genome size (Mbp)
Haemophilus influenzae	childhood meningitis	1.8
Helicobacter pylori	ulcer	1.7
Mycoplasma pneumoniae	bacterial pneumonia	0.8
Mycobacterium tuberculosis	lung tuberculosis	4.4
Treponema pallidum	syphilis	1.1
Mycobacterium leprae	leprosy	3.3

Microorganisms: isolation, preservation, safety

General. For most experiments with microorganisms, pure cultures are used. In biotechnology, most strains have additionally been optimized for a specific application, using rounds of mutation and selection. Microorganisms are maintained and conserved in culture collections. They are propagated on solid or liquid nutrient media under sterile conditions. Most microorganisms used in biotechnology grow aerobically on organic substrates (heterotrophic growth). Photosynthetic microorganisms are cultured under light, anaerobic bacteria under the exclusion of oxygen.

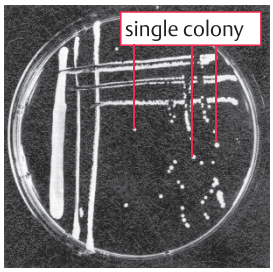
Pure cultures are obtained from culture collections or from their natural habitats (soil, water, food, other organisms) using enrichment cultures. The preferred method for obtaining a pure culture is the streak plate method, in which a mixed culture is spread over the surface of a sterile nutrient agar (a crosslinked polysaccharide isolated from marine algae) with a sterile wire loop (plating). Usually, growth conditions are chosen (→88) that favor the microorganism one wants to isolate (selection) (→24): for example, excluding oxygen and working under light with CO_2 as the sole source of carbon and N_2 as the sole nitrogen source leads to enrichment in cyanobacteria. A sugar medium at slightly acidic pH enriches fungi, incubating at elevated temperatures favors thermotolerant microorganisms, and when casein is the sole nitrogen source, protease-secreting microorganisms have a selective advantage. Based on 16S-rRNA analysis, however, it is believed that < 5 % of all naturally occurring microorganisms can be isolated by these methods (→74).

Culture collections are used to conserve pure cultures. The identity, viability, and metabolic functions of conserved cultures must be tested upon reactivation. The conventional method for conservation consists of transferring a pure culture at regular time intervals to a new agar plate or slant. This method may lead, however, to degeneration. Important type or production strains are therefore preserved under either of the following conditions: 1) under metabolically inert liquids such as mineral oil (suitable for hyphae-forming fungi); 2) freezing at –196 °C in liquid N_2 or at –70 °C in a deep-freezer; freezing and thawing must be done rapidly and in the presence of glycerol to prevent cell destruction by ice crystals (this method is mainly used for bacteria and yeasts); 3) vacuum drying of cell suspensions on a carrier (sand, silica gel) and in the presence of a mild emulsifier (skim milk, serum) and preservation at –70 °C. In all cases, it must be verified that the conserved strains can be reactivated. Most nations operate large public culture collections from which pure cultures can be ordered. They are either universal for all types of microorganisms (e. g., the American Type Culture Collection, ATCC, or CABRI, Common Access to Biological Resources and Information, a European consortium of general resource collections, e. g., the German *Deutsche Sammlung für Mikroorganismen und Zellkulturen*, DSMZ, and specialized collections for particular groups of microorganisms, such as the Dutch *Centralburau voor Schimmelkultuuren* CBS). All industrial companies that produce biotechnological products, and many hospitals, have their own culture collections. If the value of a strain lies in plasmid-coded properties (e. g., in the generation of libraries of plasmid-coded enzyme mutants), the preservation of plasmids instead of bacterial strains has become the method of choice. To this end, so-called "plasmid preps" are preserved at – 20 °C and can be stored long-term if no nucleases are present. As compared to whole strains, plasmid preps are also simpler to transport or send to other laboratories.

Safety. Each study using microorganisms must comply with biological safety rules (→332), because dangerous pathogens may occur in all microbial isolates (examples: *Bacillus subtilis:* harmless producer of technical enzymes, *Bacillus anthracis*: anthrax pathogen; *Aspergillus oryzae*: used for soy sauce production, *Aspergillus flavus*: forms highly hepatotoxic and carcinogenic aflatoxins). For safety considerations, microorganisms are classified into four risk groups. Both the construction and the equipment of a laboratory and the operating rules must be adapted to the relevant risk group. Risk group 1 (generally safe) includes microorganisms that have been used in food production for centuries, e. g., *Saccharomyces cerevisiae* and *Aspergillus oryzae*. Most microorganisms used in biotechnology fall into risk group 1.

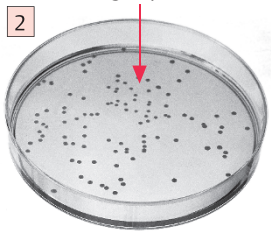
Pure cultures

1 streak plate method using nutrient agar



transfer of single colonies in liquid culture or onto nutrient agar: pure culture

2

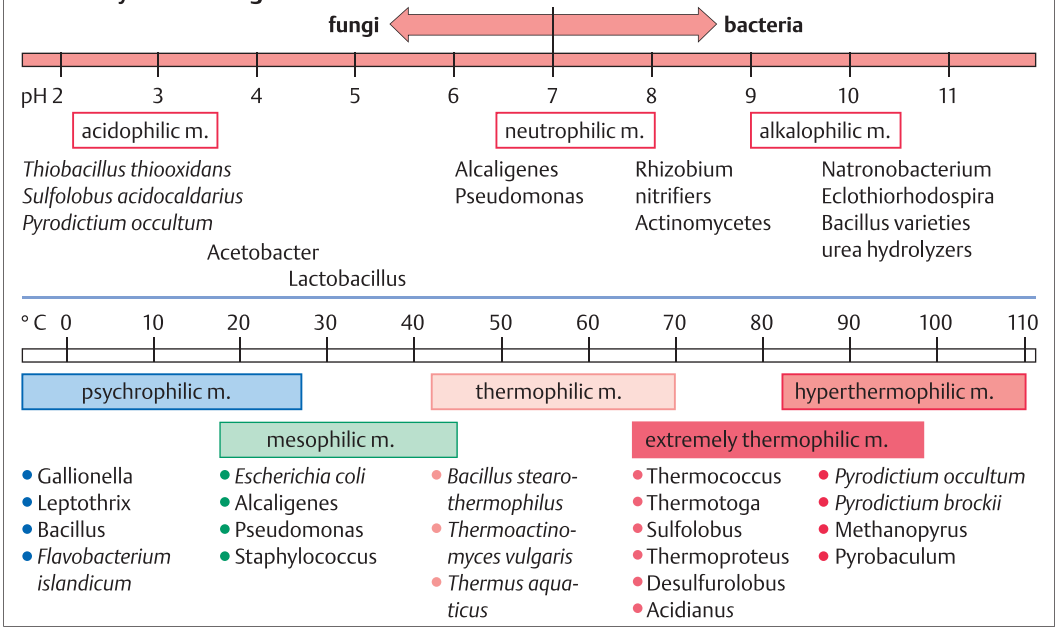


Enrichment cultures (examples)

bacteria		energy source, nutrients
phototropic		
Rhodospirilla	●	light, H_2 or organic acids, CO_2
Cyanobacteria	■	light, CO_2, N_2 as nitrogen source
chemolithotrophic		
Nitrosomonas	●	NH_4^+ as H donor, O_2 as H acceptor
Thiobacillus	●	H_2S, S or $S_2O_3^{2+}$ as H donor
methane formers	■	H_2 as H donor, CO_2 as H acceptor
heterotrophic		
Pseudomonads	■	2 % KNO_3 as H acceptor, organic acids
Clostridia	■	starch, NH_4^+, pasteurized inoculate
Enterobacteria	■	glucose, NH_4^+
lactic acid bacteria	■	glucose, yeast extract, pH 5
Bacilli	●	starch, NH_4^+
Streptomycetes	●	mannitol, NH_4^+
enzyme secretors		
protease-forming strains	●	glucose, NH_4^+, casein
lipase-forming strains	●	glucose, NH_4^+, tributyrin

● aerobic or ■ anaerobic growth conditions

Diversity of microorganisms



Risk groups (selection)

risk group 1	risk group 2	risk group 3
Acetobacter acetii, *Agrobacterium tumefaciens*, *Bacillus subtilis*, *Lactobacillus casei*	*Acinetobacter calcoaceticus*, *Escherichia coli*, *Pseudomonas aeruginosa*	*Bacillus anthracis*, *Mycobacterium tuberculosis*, *Yersinia pestis*
Penicillium notatum, *Rhizopus oryzae*, *Aspergillus niger*, *Candida tropicalis*	*Aspergillus flavus*, *Candida albicans*, *Trychophyton rubrum*, *Histoplasma capsulatum*	*Histoplasma capsulatum*

bacteria fungi, yeasts

Microorganisms: strain improvement

General. Microorganisms isolated from environmental samples rarely exhibit all the properties that are required in a technical application. Thus, they are usually optimized by a series of mutation and selection steps. The targets of strain improvement are usually: 1) to increase the yield of the desired product; 2) to remove undesired by-products; and 3) to improve general properties of the microorganism during fermentation (e. g., reduced fermentation time, no interfering pigments formed, resistance to bacteriophages). A great advantage in dealing with microorganisms is their short doubling time (often < 1 h): it allows a very large number of mutants to be produced and screened in a short time. In eukaryotic organisms, e. g., fungi, recombination events must be taken into account. With increasing knowledge of microbial metabolism, its regulation and its coding by the genome, genetic methods that delete or amplify defined metabolic steps in a targeted way are on the increase (metabolic engineering).

Mutation. The spontaneous mutation frequency (changes in DNA sequence due to natural mutation events and errors during replication) is on the order of 10^{-7} for a gene (1,000 bp) of normal stability. Most mutations remain silent or they revert genetically or functionally or by DNA repair mechanisms to the original state. Thus, for industrial strain improvement harsher mutation conditions are required: the use of UV radiation or of mutagenic chemicals are methods of choice, and, depending on the experimental goals, conditions are chosen to achieve a mortality rate of 90 % to > 99 %. Survivors exhibiting the desired properties are then selected according to their phenotypes.

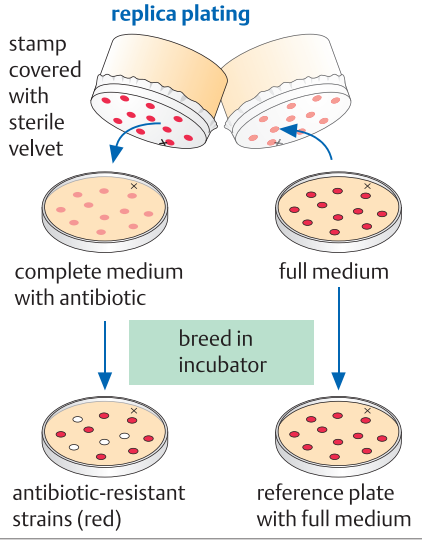
Selection using surface cultures. Phenotype selection is often synonymous with the selective isolation of mutants with high productivity. A key requirement for such experiments is the availability of an indicator reaction. For example, the resistance of a mutant to antibiotics, inhibitors, or phages can be identified if the mutant can grow on a nutrient agar that contains one of these agents. Replica plating first on a nutrient-rich agar, followed by plating on a selection medium, may yield very useful information. An enrichment step in a penicillin-containing agar (penicillin inhibits only growing cells) can help to identify auxotrophic mutants, which depend on the presence of a given metabolite for growth. If mutants that form a biologically active metabolite (e. g., an antibiotic or an enzyme) in higher yields are to be isolated, the size of inhibition or lysis plaques can be used as an indicator. Thus, if, e. g., lipases are being screened, the diameter of a halo around a clone growing on an agar plate which appears opaque due to its tributyrin content provides a first guess as to the amount of lipase produced. The great advantages of such selection procedures are 1) high flexibility in the choice of the selection criterion and 2) high number of mutants that can be visually screened (several hundred on a single agar plate). If such simple procedures are not available, a high-throughput assay must first be developed. Many procedures have been described to this end. They comprise biochemical indicator reactions, immunoassays or, in the worst case, analysis of each mutant cell, usually distributed into microtiter plates, through HPLC, capillary electrophoresis or similar procedures. Due to the random method of mutagenesis, however, the strains obtained by this kind of selection are usually defective in several genes and must be tested for their robustness as production strains in separate experiments. To this end, they are subjected to further selection with respect to growth, productivity, and other features using shake flasks and then small bioreactors under conditions resembling the production process. The best candidates may then be backcrossed with wild-type or less mutated strains to reduce the negative effects arising from many passages of random mutation.

Selection in submersed culture. Continuous fermentation has also been used for selecting microorganisms. A pure culture of a microorganism is grown in a chemostat in the presence of a mutagenic agent and subjected to selective pressure, e. g., by gradually replacing a good carbon source (→88) with a poor one. During continuous growth, those mutants that are better adapted to the altered growth conditions prevail. This method cannot be used, however, for selecting mutants that form a desired metabolite in higher concentrations.

Strain improvement of microorganisms

mutagens	mechanism	applications
physical agents		
ionizing radiation (x-ray)	leads to single- and double-strand DNA breakage	major genetic alterations
UV light (254 nm)	thymidine and cytidine form dimers	point mutations
chemical agents		
nitrite	deaminates adenine to hypoxanthine, cytidine to uridine	point mutations
alkylating agents	alkylate purines	point mutations
base analogs	are incorporated into replicated DNA	major genetic alterations
acridine orange	intercalates into DNA	major genetic alterations
biological agents		
transposons	transfer DNA elements within a chromosome	gene markers

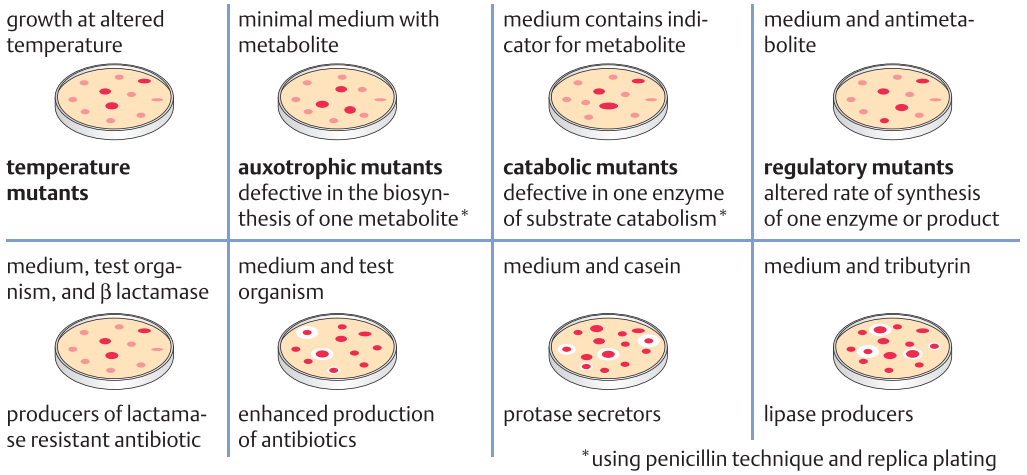
Replica plating and penicillin technique



penicillin technique

cell suspension of Ampicillin-resistant bacteria
↓ mutagenize
mixture of wild type and mutants
↓ grow on complete medium
wild type and mutants grow
↓ grow on minimal medium with Ampicillin
only wild type grows and is killed by Ampicillin
↓ remove Ampicillin, grow on complete medium
only mutants grow
↓ replica plating on minimal and complete medium
auxotrophic mutants grow only on complete medium

Selection media



*using penicillin technique and replica plating

Biochemistry

General. Biochemistry deals with the chemistry of life. It describes the building blocks of living cells, their synthesis and degradation, their overall metabolism and its regulation, and the thermodynamic rules involved in these processes (bioenergetics). Biochemistry is the cornerstone of molecular genetics and cell biology, and biochemical knowledge is essential for the understanding of biotechnology and genetic engineering. The following pages will provide a brief introduction but cannot substitute for a thorough study based on textbooks.

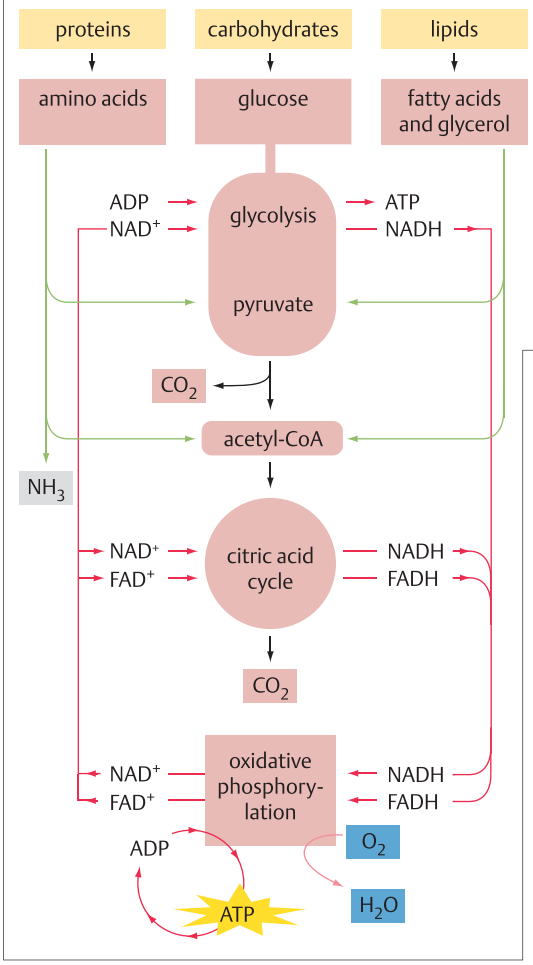
Building blocks. The essential chemical building blocks of life are the amino acids (→124) and the peptides and proteins derived from them, sugars and the oligo-, polysaccharides and glycosides which they form, fatty acids and their esters (triglycerides, phosphoglycerides) (→162), and nucleosides (→136) and their derivatives (of which nucleic acids and DNA is one group). Many of these building blocks are chiral, i. e., only a selection of all possible enantiomers is being used. Thus, all chiral proteinogenic amino acids are of L chirality. The biosynthesis, conversion and degradation of these building blocks proceeds *via* catalytically active proteins, the enzymes (→166). Many of the above compounds are used for technical purposes and thus are target compounds of biotechnology.

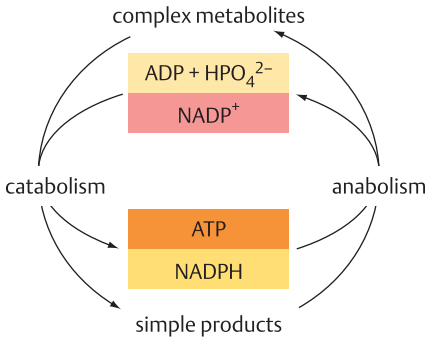
Metabolism. (→36) The original building blocks of all biochemicals are CO_2, inorganic nitrogen, inorganic phosphorous, a few minerals such as iron ions, and water. A comprehensive biosynthesis of biochemicals from these base materials is mostly limited, however, to plants or autotrophic microorganisms that assimilate CO_2 using solar energy (photosynthesis) and water into carbohydrates. Most living organisms do not possess this capacity and thus form their metabolites from organic nutrients through biotransformation. The buildup of cellular building blocks is termed biosynthesis or anabolism, and their degradation (and the dissection of food into metabolically useful compounds) is called biodegradation or catabolism. Though the number of known metabolites exceeds 50,000, there are only a few thousand essential metabolites that are formed and transformed through central metabolic pathways which are quite similar throughout living organisms. Such metabolic modules are, e. g., the biodegradation of glucose to activated acetic acid (acetyl-CoA)(glycolysis), the formation of C-6 compounds from acetyl-CoA and oxalacetate (tricarboxylic acid cycle) or the transformation of fatty acids to glucose (gluconeogenesis). This observation and the universality of the genetic code have led to the modern discipline of a "synthetic biology:" (→320) it is now possible to create artificial metabolic pathways, to express and optimize them *via* "metabolic engineering" in suitable host organisms (plants, microorganisms), and to apply this technology to the synthesis of value-added chemicals. A special challenge for using this technology in eukaryotic organisms resides in the fact that many metabolic and bioenergetic functions are partitioned among different subcellular compartments; thus, glycolysis occurs in the cytoplasm, the citric acid cycle and oxidative phosphorylation of eukaryotic cells resides in their mitochondria, and the degradation of cellular components in the lysosomes.

Metabolic regulation. The spatial, temporal and environment-dependant regulation of metabolism proceeds on the level of molecular genetics (induction, gene expression, repression) as well as through the control of enzyme activity by their metabolites. Thus, many key enzymes are inhibited by the products they form (product inhibition), and enzymes at metabolic branches can be induced or inhibited by several metabolites. This type of regulation proceeds through conformational changes after binding of a metabolite (allosteric regulation). In multicellular organisms, an additional regulatory mechanism is through the formation of messenger compounds that bind to receptors in the membranes of target cells, initiating regulatory signal cascades within those cells.

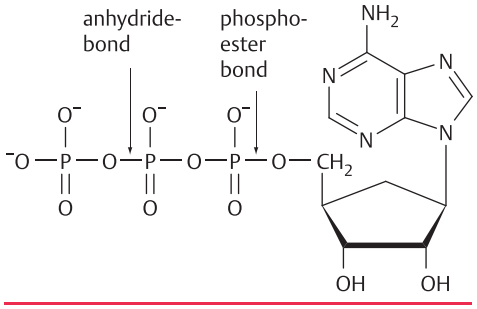
Bioenergetics. Anabolic reactions require energy, during catabolism, energy is usually generated. The "currency" for the energy requirements of a cell are energy-rich phosphate-esters such as adenosine triphosphate (ATP). The hydrolysis of ATP provides energy and thus facilitates energy-consuming (endergonic) reaction steps such as the transfer of a phosphate on a sugar molecule, which activates the sugar for further reactions. ATP is primarily generated from ADP and phosphate by oxidative phosphorylation in aerobic organisms, and by photophosphorylation in plants; the chemical energy for this highly endergonic reaction stems from a proton gradient at the cell membrane. Such transfers of energy-rich groups determine the whole metabolism and render important metabolic steps irreversible. Redox processes in metabolism occur stepwise through cofactors of different redox potentials.

Metabolites





High-energy compounds

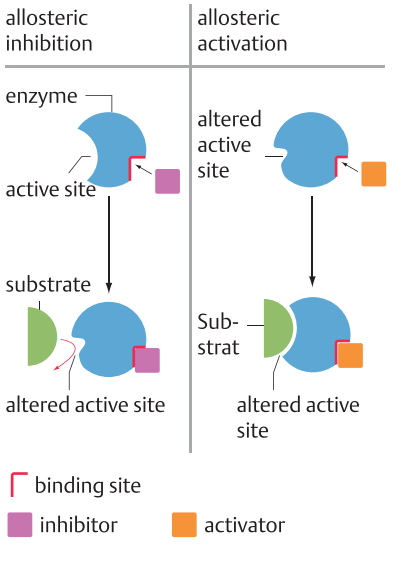


ATP ADP AMP adenosine

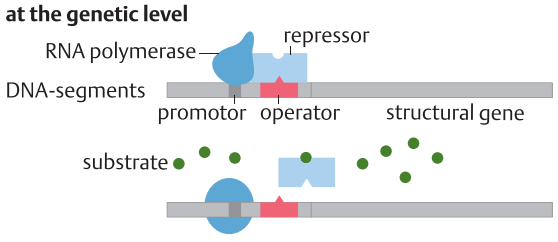
compound	free enthalpy $\Delta G^{0'}$
phosphenolpyruvate	–61,9 (kJ • mol^{-1})
ATP (→ AMP + PP_i)	–45,6
ATP (→ ADP + P_i)	–30,5
ADP	–19,2
glucose-6-phospate	–13,8

Regulation

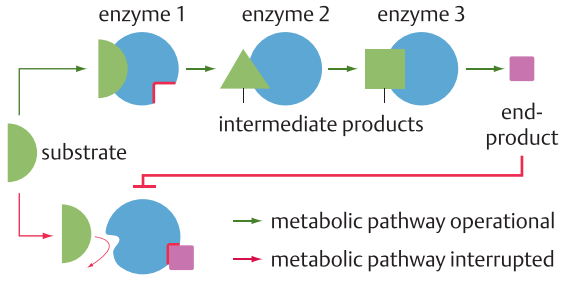
at the enzyme level



at the genetic level



at the metabolic level



Amino acids, peptides, proteins

Amino acids. Nearly all peptides and proteins are built from just 20 different amino acids (→124); 19 out of them are chiral ("optically active") with L-configuration. Amino acids may form peptide bonds, thus condensing to peptides or folding to proteins. The spatial configuration of peptides and proteins is determined by the sequence of the amino acids and the properties of their side chains. These can be polar (hydroxy-, amide- or thiol-groups) or charged (carboxy- or amino groups), thus allowing for the formation of intermolecular ionic bonds or hydrogen bridges, or they can be of low polarity (alkyl groups, aromatic residues, secondary amides, thiol ether) thus leading to intermolecular hydrophobic interactions. All amino acids have at least two, sometimes three ionizable groups, which are charged or uncharged depending on the solution pH. At their isoelectric point, or pI, they have no net electric charge, i. e., the amino acid does not migrate during electrophoresis. Peptides and proteins with their many ionizable groups also show a pI. Some peptides and proteins contain special amino acids and amino acid derivatives, other than the 20 "proteinogenic" amino acids. Thus, collagen fibers contain 4-hydroxyprolin and 5-hydroxylysine. Specific non-proteinogenic amino acids such as γ-aminobutyric acid (GABA) or degradation products such as histamine are neurotransmitters. Peptide antibiotics often contain unusual amino acids. Several amino acids are produced in large quantities, e. g., as food and feed additives.

Peptides are often used as signal molecules. They form antioxidants such as glutathione, hormones such as insulin (→222) and growth factors (→224) such as the granulocyte-CSF (colony-stimulating factor) (→238). Peptides are formed from amino acids via planar *cis*-peptide bonds, which have limited rotational freedom thus largely reducing the number of conformations of a peptide chain. The permitted conformations of a peptide are represented in a Ramachandran diagram. Peptides form secondary structures: due to the number of intermolecular hydrogen bridges between N-H and C=O groups, a helical polypeptide conformation with 3.6 amino acids per turn is particularly stable (α-helix). The longer stretched 3_{10} helix, by comparison, occurs only in some 10 % of all protein helices. The same type of hydrogen bridge between two neighboring polypeptide chains results in ß-fold structures. The topology of a peptide or protein (tertiary structure) is a consequence of the sequence and side-chain structure of the constituting amino acids.

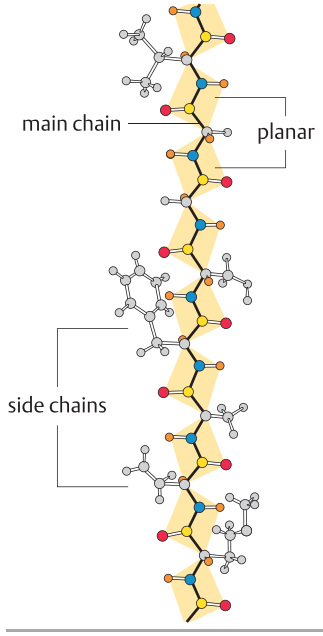
Proteins consist of one or several peptide chains that form a unique structure, mostly by non-covalent interactions. Globular proteins consist of ~31 % α-helix- and ~28 % ß-folds. The remaining ~40 % of the sequence form *random coil*s and turns. ß-turns are composed of 4 amino acids with glycine in position 3. They often connect two ß-fold structures and initiate a directional turn of the peptide chain. Using such partial structures, bioinformatic programs can often predict partial structures or even structures of proteins with significant precision. Protein structures are often termed as primary structures (= amino acid sequence), secondary structures (helices and ß-turns), tertiary structures (3D structure of a single polypeptide chain) and quaternary structures (3D structures if several peptide chains form a protein). By 2014, the 3D structures of about 100.000 proteins had been solved, mostly through x-ray analysis of protein crystals, in several cases also by high-resolution NMR spectroscopy of dissolved proteins (→324). Using this large knowledge base, predictions about the structure of a protein can today often be based on its amino acid sequence, using bioinformatic methods. This is of particular importance in the case of membrane proteins which usually are hard to crystallize and thus cannot be subjected to x-ray analysis. Recently, however, diffraction of very small and thin crystals has also become possible by electron beam analysis in a synchrotron. The size of a protein is given in kilo-Dalton (kDa), its sequence by a three- or one-letter code for each amino acid. Proteins are often decorated with other groups, e. g., with heme or phosphate groups, metal ions or sugar side chains. Apart from being the catalysts of metabolism, specialized proteins have many important functions in an organism. Thus, serum albumen, hemoglobin and ferritin are blood proteins (→226); proteins of the immune system regulate and coordinate defense mechanisms or the growth of specific cell types. And proteins such as myosin (muscle) or collagen (connective tissue) participate in forming the shape of an organism.

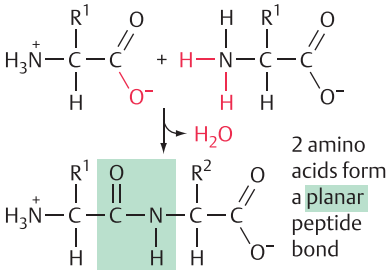
The protein-forming ("proteinogenic") amino acids

amino acid	3-letter code	1-letter code	type	pi
alanine	Ala	A	hydrophobic	6.00
arginine	Arg	R	basic	11.5
asparagine	Asn	N	polar	5.41
aspartic acid	Asp	D	sour	2.77
cysteine	Cys	C	polar	5.02
glutamine	Gln	Q	polar	5.65
glutamic acid	Glu	E	sour	3.22
glycine	Gly	G	polar	5.97
histidine	His	H	basic	7.47
isoleucine	Ile	I	hydrophobic	5.94
leucine	Leu	L	hydrophobic	5.98
lysine	Lys	K	basic	9.59
methionine	Met	M	nonpolar	5.74
phenylalanine	Phe	F	hydrophobic	5.48
proline	Pro	P	hydrophobic	6.30
serine	Ser	S	polar	5.68
threonine	Thr	T	polar	5.64
tryptophan	Trp	W	hydrophobic	5.89
tyrosine	Tyr	Y	polar	5.66
valine	Val	V	hydrophobic	5.96

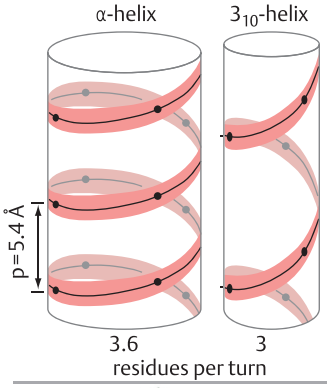
pi = isoelectric point

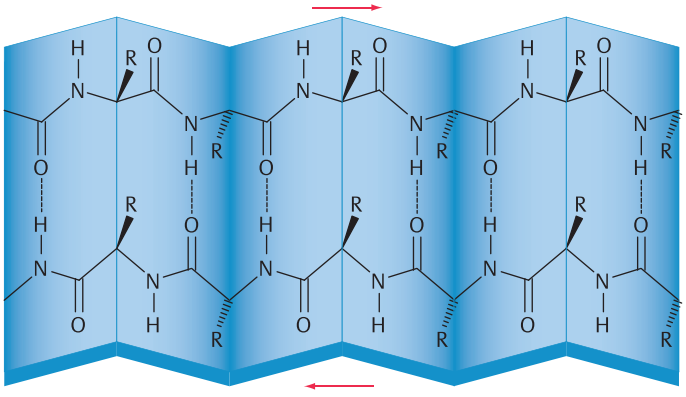
Peptide bond/polypeptides

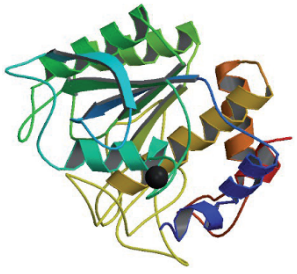




α-helix and β-pleated sheet







Subtilisin Carlsberg, a bacillus protease (27.3 kDa, 1 peptide chain, 274 amino acids)
ribbons: α-helix
arrows: β pleated sheet.
Protein chains are colored from the N-terminal to the C-terminal using a rainbow (spectral) color gradient
kDa =kilo-Dalton; 1 Dalton is 1/12 of the mass of a ^{12}C carbon isotope

Enzymes: structure, function, kinetics

General. Enzymes are proteins with catalytic activity. They transform substrates into products, often with a high velocity of >1000 turnovers/sec. Some enzyme require cofactors, such as NADH/NADPH, FADH, pyridoxal phosphate, ATP, heme or metal ions, for their catalysis. Enzymes are the catalytic agents of biotransformations and are usually highly regio- and stereoselective. Their metabolic reaction products are termed metabolites. Their activity is regulated on the genetic level by induction or repression, and on the enzyme level through interaction with metabolites (product inhibition, allosteric control). Many enzymes are already used in technical processes (enzyme technology) (→168). Their advantage lies in their selectivity, but also in their high catalytic activity under ambient conditions.

Nomenclature. According to international conventions, enzymes are classified by 4-digit EC-numbers into six groups according to their reaction type (hydrolysis, redox reactions etc.) (→166). Presently (2014), more than 6,500 enzymes have been classified.

Enzyme catalysis. Enzymes accelerate chemical reactions by lowering the activation energy. This can be achieved by a combination of several mechanisms: a) through acid-base catalysis, b) through covalent intermediates, c) through metal-ion catalysis, d) through electro-static catalysis, e) with the aid of neighboring groups and orientation effects, and f) by preferred binding of a transition state. For example, serine proteases such as subtilisin Carlsberg (→176) lower the activation energy for hydrolysis through three factors: 1) the formation of an "active site" which binds the peptide substrate specifically through steric and electrostatic interactions, 2) the activation of the carbonyl group of the peptide bond to be cleaved with the hydroxyl group of a catalytically active serine, forming an energetically favored tetrahedral transition state which is stabilized by hydrogen bonds among a neighboring aspartic acid and a histidine side chain ("catalytic triad"), and 3) the location of this reaction within an environment of hydrophobic amino acid side-chains, i. e., in an anhydrous milieu. Our knowledge about enzyme catalysis is largely based on the x-ray structure analysis of enzymes in the presence of substrate analogues, on comparative studies using genetically engineered enzyme variants, and on chemical model reactions.

Enzyme kinetics. Enzyme reactions can be described by chemical reaction equations. Investigation of the kinetics involved leads to valuable information about the mechanism of the particular enzyme, and its inhibition (e. g., competitive or non-competitive inhibition). Many enzyme reactions follow approximately the Michaelis-Menten equation, which is based on the measurement of enzyme reactions of zero order (velocity is independent from substrate concentrations) according to the reaction

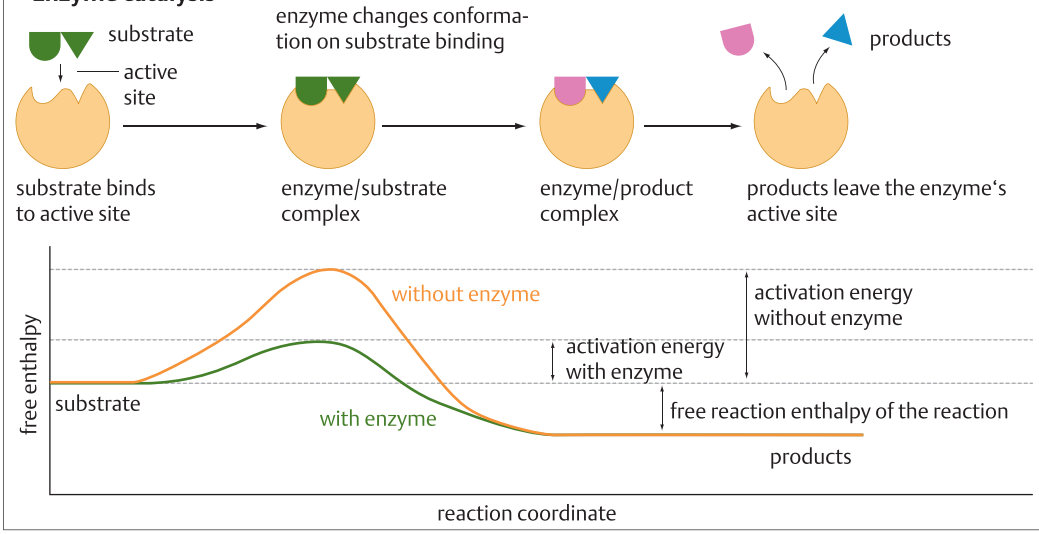
Enzyme (E) + Substrate (S) ⇋ Enzyme-Substrate or Michaelis-Complex (ES) ⇋ Product (P) + Enzyme (E), with the equilibrium assumption $dES/dt = 0$

The reaction process is usually represented in a diagram that shows the initial enzyme velocity (which can be easily determined) as a function of substrate concentration. The maximal reaction velocity, V_{max}, is reached when the enzyme is saturated with its substrate, i. e., when all the enzyme exists as an enzyme-substrate complex. K_M, the Michaelis constant, describes the substrate concentration at which the reaction velocity reaches 50 % of the maximum. K_M and V_{max} can be determined graphically using the Lineweaver-Burk diagram where substrate concentration and reaction velocity are plotted in a double reciprocal manner. This leads to a linear relationship where the intersection with the y-axis is $1/V_{max}$ and with the x-axis is $-1/K_M$. From these values, both V_{max} and K_M can be obtained by simple calculations. The turnover number k_{cat} of an enzyme is given by

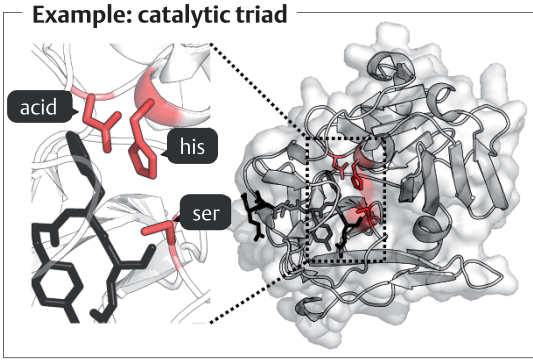
$$k_{cat} = V_{max}/E_T$$

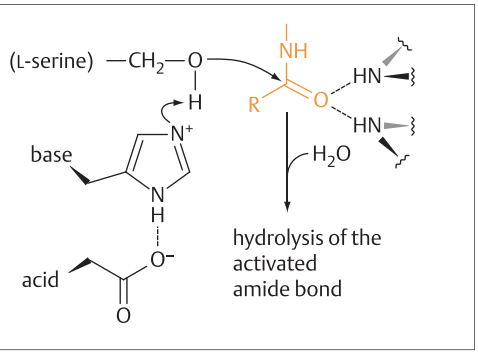
and k_{cat}/ K_M is a measure for the catalytic efficiency of an enzyme (E_T = total enzyme present). Some enzymes exhibit a turnover number that approach the limits of diffusion-controlled reactions: each collision of an enzyme with its substrate leads to a reaction. Different from the assumptions made by the Michaelis-Menten calculation, 60 % of all enzyme reactions occur with two substrates (in the case of two substrates and two products: "bi-bi reactions"). This leads to much more complex kinetic equations. The knowledge of enzyme kinetics is indispensable for the engineering of an enzyme reactor (→102) in view of overcoming inhibitor actions.

Enzyme catalysis

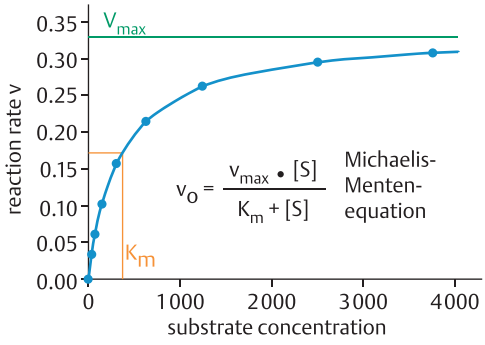


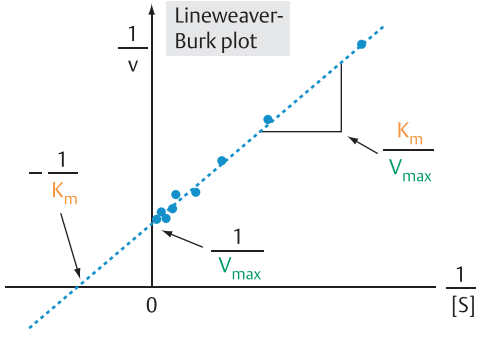
Example: catalytic triad





Enzyme kinetics





Kinetic data of some enzymes

enzyme	EC-number	K_M (M)	k_{cat} (s^{-1})**	k_{cat}/K_M ($M^{-1}\ s^{-1}$)
acetylcholinesterase	EC 3.1.1.7	9.5×10^{-5}	1.4×10^{4}	1.5×10^{8}
carbonate dehydratase	EC 4.2.1.1	1.2×10^{-2}	1.0×10^{6}	8.3×10^{7}
catalase	EC 1.11.1.6	2.5×10^{-2}	1.0×10^{7}	4.0×10^{8}
chymotrypsin*	EC 3.4.4.5	6.6×10^{-4}	1.9×10^{2}	2.9×10^{5}
fumarate hydratase	EC 4.2.1.2	5.0×10^{-6}	8.0×10^{2}	1.6×10^{8}
superoxide dismutase	EC 1.15.1.1	3.6×10^{-4}	1.0×10^{6}	2.8×10^{9}
urease	EC 3.5.1.5	2.5×10^{-2}	1.0×10^{4}	4.0×10^{5}

*with N-acetyltyrosine ethyl ester as substrate **turnover number

Sugars, glycosides, oligo- and polysaccharides

General. The building units of all sugars are the monosaccharides. They are formed in the metabolism of heterotrophic organisms by gluconeogenesis, and by photosynthesis from CO_2 and H_2O ("carbohydrates" in phototrophic organisms such as plants. A key compound is D-Glucose, a hexose, but other sugars also play important roles in metabolism. Thus, ribose and deoxyribose, two pentoses, are among the main components of nucleic acids. Oligosaccharides are composed of a few sugar units while polysaccharides are composed of many. Mono- and oligosaccharides are often linked to proteins (glycoproteins) or lipids (lipoproteins); they are often termed "glycoconjugates" and often display structural or regulatory functions. Many enzymes produced by biotechnology are glycoproteins. The generation of "humanized" glycoconjugates in a host organism used for biotechnology is important, but no simple task (glycobiology) (→262).

Monosaccharides are aldehyde- (aldoses) or keto-derivatives (ketoses) of straight-chain polyhydroxy alcohols. Sugars of chain-length 3 (C-3) have one stereocenter and thus can form 2^1=2 stereoisomers. Their absolute configuration depends on their simplest representative, D- and L-glyceraldehyde. C-4 sugars have two stereo-centers, forming four stereoisomers, C-5-sugars have three (eight stereoisomers) and C-6-sugars have four (16 stereoisomers). Only a few of these stereoisomers are central products of metabolism, but most of all the possible stereoisomers were already found in natural compounds such as antibiotics (rare sugars). C-3 and C-4 sugars occur as open chains, C-5 and C-6-sugars form cyclic half acetals or half ketals (C-5: furanoses, C-6: pyranoses); in this reaction, a new stereocenter is formed at the acetal/ketal-carbon (anomeric carbon), which is usually termed an α- or ß- glycosidic bond. Upon binding of an alcohol to an anomeric carbon atom, acetals or ketals are formed. Sugars are mostly drawn in the form of planar Haworth-formula, but their tetrahedral sp^3-carbons lead to chair- or boat-conformations where the polar hydroxy groups are positioned in a staggered, preferentially eclipsed conformation. The preferred conformation of a sugar can be deduced using NMR spectroscopy. D-Glucose in aqueous solution occurs nearly exclusively as ß-D-glucopyranose, D-ribose exists in an equilibrium of 75% ß-D-ribopyranose and 25% ß-D-ribofuranose.

Sugar derivatives. During metabolism, monosaccharides are frequently oxidized at position C1 to aldonic acids; in this reaction, D-glucose forms D-glucuronic acid. Vitamin C (L-ascorbic acid) (→134) is a partially oxidized pentose that forms an internal ester bond. Reduction of the aldehyde- or keto group leads to alditols. Important lipid components such as glycerol (C-3) and *myo*-inositol fall in this group. Reduction of one hydroxyl group leads to deoxysugars such as 6-deoxy-L-Mannose (α-L-Rhamnose), a component of pectin (→186). Single hydroxyl groups can be substituted by amino groups such as in N-Acetyl-mannosamine. By aldol condensation with pyruvate, N-acetyl neuraminic acid is formed, a component of glycoproteins and bacterial cell walls. Biochemically important compounds are the phosphate esters of sugars. As energetically "activated sugars," they play a key role in metabolism, in biosynthesis and in intercellular signaling (→76).

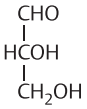
Oligosaccharides are formed through glycosidic linkages and cleaved by glycosidases. The formation of a glycosidic linkage proceeds via activated sugars, usually through nucleotide sugars such as uridine-diphosphate-glucose, and leads to the synthesis of optically active glycosides that are termed α- or ß, as in the case of acetals and ketals. The simplest oligosaccharides are disaccharides such as sucrose or lactose (→116, 188).

Polysaccharides (glycans) (→158) are natural compounds occurring, e. g., as structural components of plants and insects (cellulose, pectin, chitin), as fibers (cotton) or as energy storage (starch). They are formed from activated monosaccharides through glycosidic linkages. They are classified into homo- and heteropolysaccharides, and, depending on their building blocks, in glucanes (built from glucose residues) or galactans (built from galactose). Polysaccharides can be unbranched, like cellulose, or branched, like starch. They are degraded by glycosidases. In biotechnology, starch (→176), cellulose and hemicelluloses (→182) are C-sources for fermentation, and animal, plant or microbial polysaccharides such as hyaluronic acid, alginates or dextran are used in medical treatment or as fine chemicals. Microbial heteropolysaccharides such as xanthan are used as thickeners in food and tertiary oil recovery.

Monosaccharides

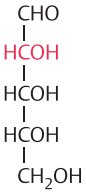
aldoses

C-3



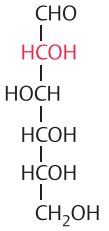
D-glycer-aldehyde

C-5

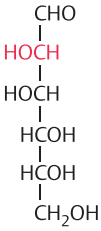


D-ribose (Rib)

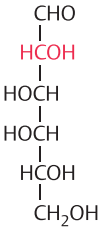
C-6



D-glucose (Glc)



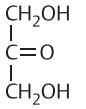
D-mannose (Man)



D-galactose (Gal)

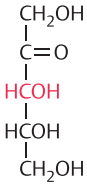
ketoses

C-3

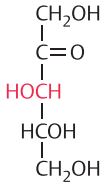


dihydroxy-acetone

C-5

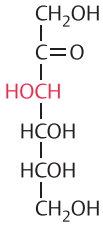


D-ribulose



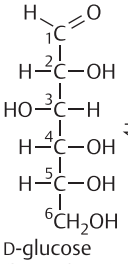
D-xylulose

C-6

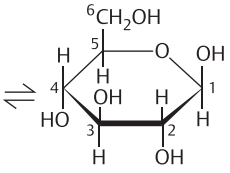


D-fructose

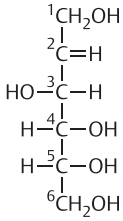
Half acetals and glucosides



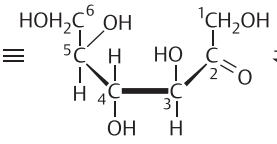
D-glucose (linear form)

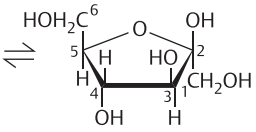


β-D-glucose (Haworth projection)

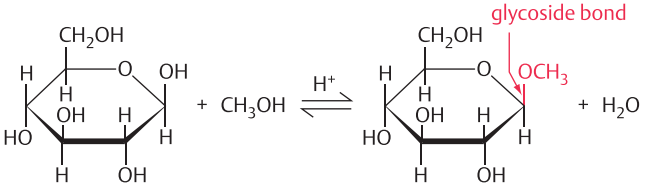


D-fructose (linear form)





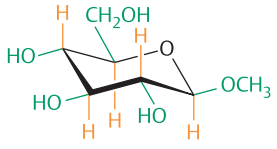
β-D-fructose (Haworth projection)



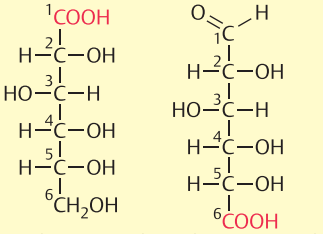
D-glucose

methyl-β-D-glucoside

Chair form

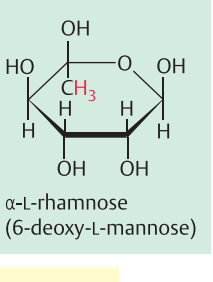


Sugar derivatives

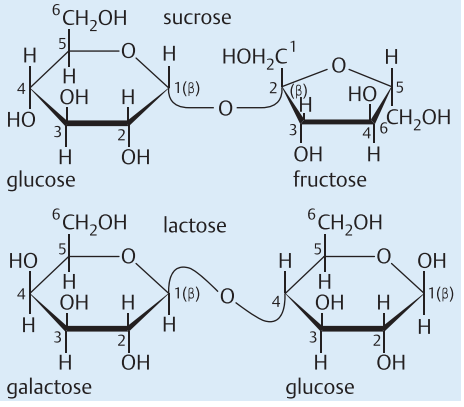


D-gluconic acid

D-glucuronic acid



α-L-rhamnose (6-deoxy-L-mannose)



sucrose

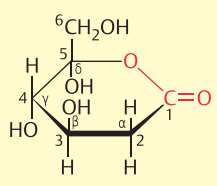
glucose

fructose

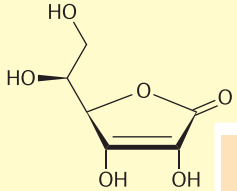
lactose

galactose

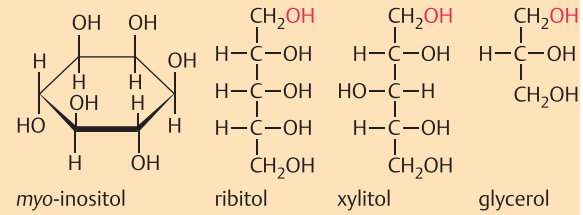
glucose



D-glucono-δ-lactone



L-ascorbic acid



myo-inositol

ribitol

xylitol

glycerol

acidic sugars
deoxy sugars
disaccharides
sugar alcohols

Lipids, membranes, membrane proteins

General. Lipids are classified into triglycerides (→162), phospholipids, sphingolipids and sterols. In aqueous solution, lipids aggregate and form micelles, double layers and membranes. The latter enclose most living cells, but may also form intracellular compartments such as organelles. Sterols are components of many biological membranes and regulate membrane fluidity. In higher organisms, some steroids function as hormones (→252). Lipids and lipoproteins (aggregates of lipids and proteins which are not covalently linked) participate in many biological transport and signaling processes.

Triglycerides (fats and oils) (→162) are esters of glycerol with fatty acids and serve mostly for energy storage. Most fatty acids have a chain length between C-12 and C-18. They can be saturated or unsaturated, with one or several double bonds. Their melting point decreases with the proportion of unsaturated fatty acids. Triglycerides are renewable raw materials and play in important role in the development of a sustainable economy (bioeconomy) (→328).

Phosphoglycerides are di-esters of *sn*-glycerol-3-phosphate with fatty acids (phosphatidinic acids); the phosphate group is esterified with glycerol and an alcohol or amine. They are amphiphilic compounds (surfactants), as their structure is composed of a polar head group and hydrophobic residues (acyl chains). They form micellar structures and membranes.

Micelles and liposomes. In aqueous solution, fatty acids form micelles above a threshold concentration (CMC = critical micelle concentration). The voluminous structure of the two acyl chains in phosphoglycerides prohibits the formation of such micelle superstructures. Instead, double layers are formed which can rearrange into liposomes, e. g. upon treatment with ultrasound. Liposomes are droplets (vesicles) of some 100Å diameter whose double membrane layer of about 60Å thickness encloses an aqueous core. Membranes and liposomes are polar at their insides and outsides but hydrophobic in their center. This architecture is the structural basis of all membranes of biological cells.

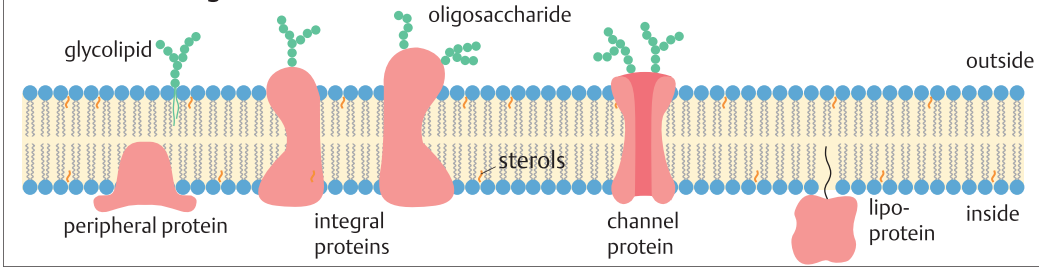
Biological membranes determine the "inside" and "outside" of cells, and also of most organelles inside eukaryotic cells (mitochondria, plastids, peroxisomes etc.). Their lipophilic components (e. g., phospholipids) diffuse very slowly between the two layers of the double membrane ("flip-flop", timescale: days), but very fast in lateral direction (timescale: minutes). Since the headgroups of phospholipids carry different charges, local "islands" of a distinct charge may develop, which helps for, e. g., the localized insertion of lipoproteins such as porins, receptors or glycolipids into a membrane. Membranes exhibit fluidity: they show a transition from an unstructured, fluid state (high lateral mobility of the phospholipids) into a more highly structured, gel-like state. The transition temperature for the fluid-to-gel state depends upon the chemical structure of the fatty acids (chain length, double bonds). Membranes of most eukaryotic cells have transition temperatures lower than their body temperature and thus are fluid. Their fluidity is often regulated by sterols. Bacteria and cold-blooded animals such as fish preserve the fluidity of their membranes in a different manner: they change the composition of their membranes in dependence of the ambient temperature through degradation and resynthesis of appropriate membrane phospholipids, thus readjusting membrane fluidity.

Membrane proteins have many and specific tasks in the uptake and secretion of substances, as well as in the communication of cells. They are classified according to integral and membrane-bound proteins. Integral proteins are identified by their amino acid sequence through longer intercepts of hydrophobic amino acids which form α-helix or ß-fold antiparallel secondary structures. Membrane-bound proteins have different structures: they contain anchor groups such as isoprenoids, fatty acids or glycosylphosphatidyl inositol (GPI), which are post-translationally linked to the peptide chain. Important groups of integral membrane proteins are the porins of the Gram-negative bacteria, the transporter proteins, the mitochondria and chloroplasts, the photosynthetic protein (photoreaction center) of plants and the visual pigment rhodopsin of animals. With connexins as the gap junction proteins between two membranes, two or more cells form connections that allow for the exchange of signals and small molecules. Examples for membrane-bound proteins are the receptors; they either function as ion channels upon binding of a ligand or as signal transducers through a downstream reaction cascade that modulates the behavior of a target cell. Membrane-bound glycoproteins and glycolipids exert important function in the immune system.

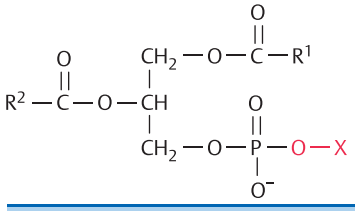
The most frequent natural lipids

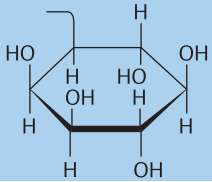
symbol	trivial name	common name	m.p. (°C)
saturated fatty acids			
12:0	lauric acid	dodecanoic acid	44.2
14:0	myristic acid	tetradecanoic acid	52
16:0	palmitic acid	hexadecanoic acid	63.1
18:0	stearic acid	octadecanoic acid	69.6
20:0	arachic acid	eicosanoic acid	75.4
unsaturated fatty acids (all double bonds are *cis*)			
16:1n–7	palmitoleic acid	9-hexadecenoic acid	–0.5
18:1n–9	oleic acid	9-octadecenoic acid	–0.5
18:2n–6	linolenic acid	octadecane-9,12-dienoic acid	–9
18:3n–3	α-linoleic acid	octadecane-9,12,15-dienoic acid	–17
18:3n–6	γ-linoleic acid	octadecane-6,9,12-dienoic acid	–10
20:4n–4	arachidonic acid	eicosa-5,8,11,14-tetraenoic acid	–49.5
20:5n–3	EPA	eicosa-5,8,11,14,17-pentanoic acid	–49.5
20:6n–3	DHA	docosa-4,7,10,13,16,19-hexaenoic acid	–44

Model of a biological membrane

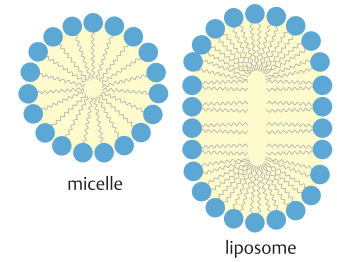


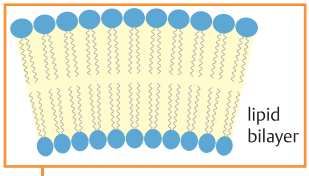
Glycerophospholipids

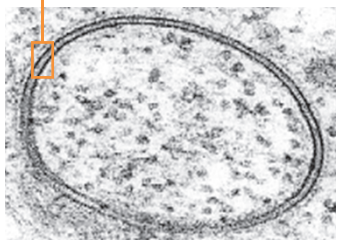


residue O—X	formula of —X	name of phospholipid
	—H	phosphatidic acid
ethanol-amine	$-CH_2CH_2NH_3^+$	phosphatidyl ethanolamine
choline	$-CH_2CH_2N(CH_3)_3^+$	phosphatidyl choline
serine	$-CH_2CH(NH_3^+)COO^-$	phosphatidyl serine
myo-inositol		phosphatidyl inositol
glycerol	$-CH_2CH(OH)CH_2OH$	phosphatidyl glycerol

Micelles and liposomes







Metabolism

General. In spite of >4 billion years of life on earth, which have led to a great variety of living organisms, their basic building blocks and their patterns of metabolism and replication are based on variations of just a few basic principles. The number of species by far exceeds 1 million, but only some 6000 different enzyme functions have been classified, and a million proteins (most of them originating from post-translational modification of a single peptide, and many of them highly homologous to simple unicellular eukaryotes such as *Saccharomyces cerevisiae*) are enough to build and maintain a higher organism such as humans. All living beings on earth form a sensitive ecological network, which includes many thousands of specialized species that survive under nearly all imaginable environmental conditions (ecological niches) (→12). A major distinction is made between autotrophic organisms, which can use CO_2 as the main carbon source, and heterotrophic organisms, which need organic compounds for growth. Another distinction is made between aerobic organisms, growing in air, and anaerobic organisms. With respect to the details of their metabolisms, organisms may differ, for example, by how they transform glucose: through fructose-1,6-bisphosphate (glycolysis) or by a pentose phosphate or a 2-keto-3-deoxy-6-phospho-gluconate pathway. With the accomplishment of total genome sequencing, a breathtaking view into even finer structural and metabolic variations has become possible, which helps us to better understand how organisms are adapted to their specific environments. We are also making great strides in understanding regulatory networks in organisms and their environmental interactions (system dynamics, biocybernetics) and are learning to simulate complex living systems *in silico* (→326). In biotechnology, a key interest is to modulate metabolism, e. g., for increased yields of a product, for elimination of a by-product, or, in breeding, for introduction or elimination of a phenotypic trait. Traditional methods of crossing or mutation, followed by selection, are now backed up more and more by genetic engineering. A new technology is to combine metabolic steps from different organisms into a new synthetic pathway, using a (mostly microbial) host organism (synthetic biology) (→320).

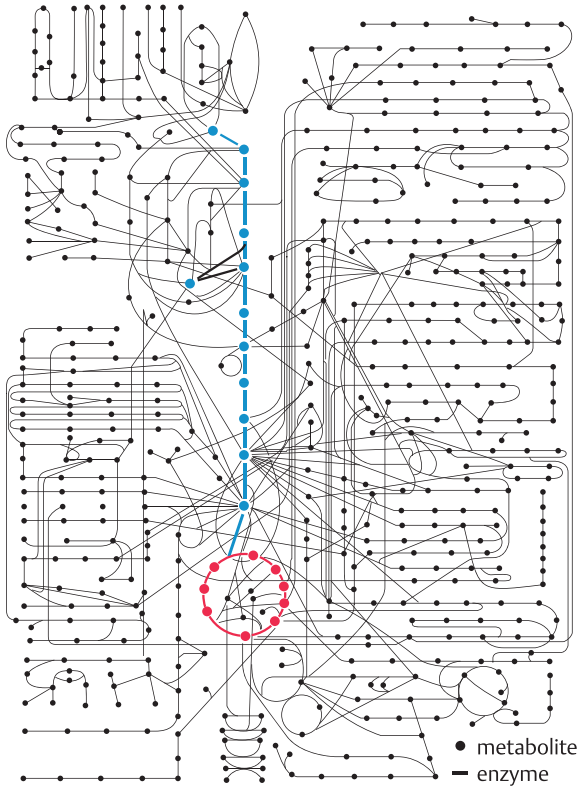
Autotrophic metabolism. Autotrophic organisms reduce CO_2 to carbon sources such as glucose. Phototrophic organisms such as plants, algae, and cyanobacteria derive the energy required for this process from light, which they convert, in their photoreaction centers, into chemical energy stored as adenosine triphosphate (ATP, "the universal energy currency"). Lithotrophic organisms derive energy from the oxidization of inorganic compounds, e. g., S, N, metal ions. Autotrophic organisms of importance in biotechnology are transgenic plants, nitrifying bacteria, and Thiobacilli used in mineral leaching.

Heterotrophic anaerobic metabolism. Heterotrophic anaerobic organisms are used in biotechnology in the production of ethanol (→138), acetone, 1-butanol (→140), and lactic acid (→148). They generate ATP by catabolism of sugars. Methane-forming Archaea, which develop in anaerobic sludge treatment (→288), also belong in this group; they exhibit some unusual metabolic steps. Energy generation in anaerobic metabolism proceeds with low yield efficiency.

Heterotrophic aerobic metabolism. Most microorganisms used in biotechnology exhibit heterotrophic aerobic metabolism. The bulk of their energy is generated through the respiratory chain, which feeds off the citric acid cycle. With ca. 36 moles of ATP generated per mole of glucose, the energy yield is very high. Several biotechnological products, such as citric acid (→146) or glutamic acid (→126), are either components of the citric acid cycle or derived from oxaloacetate or succinic acid, metabolites which must be replenished by anaplerotic pathways in overproducing organisms in order to maintain basic metabolism.

Secondary metabolism. Many organisms form metabolites that are not involved in primary cell functions (secondary metabolites). In plants, secondary metabolites are of key importance in defense mechanisms against pathogens and predators and in attracting insects for fertilization and dissemination. In microorganisms, the physiological function of secondary metabolites is less clear. Often, these are important biotechnological products such as antibiotics (→250), alkaloids, colorants, or aroma compounds.

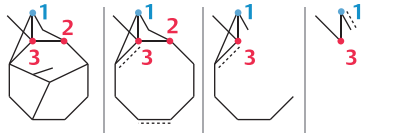
Metabolic pathways – a network



Types of metabolism

	autotrophic	heterotrophic
anaerobic	microbial leaching	ethanol acetone, 1-butanol, anaerobic sludge digestion
aerobic	transgenic plants nitrification	citric acid amino acids antibiotics enzymes vitamins, etc.

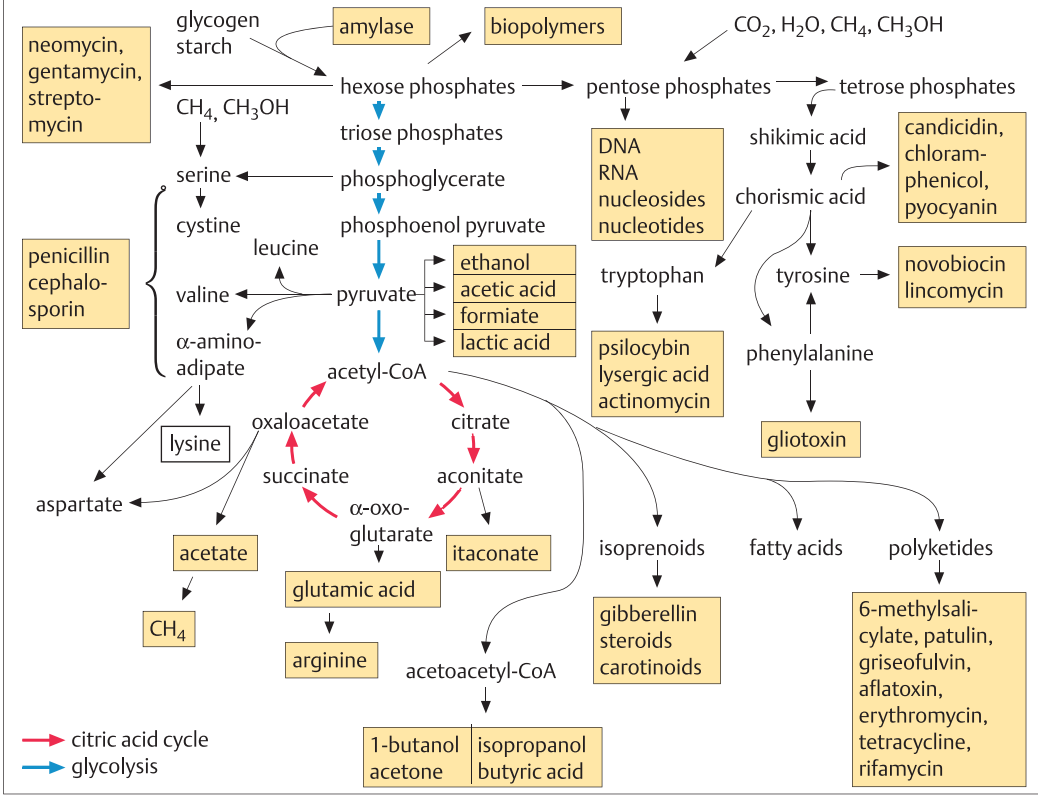
Variants of the citric acid cycle



Saccharomyces cerevisiae | *Bacillus subtilis* | *Haemophilus influenzae* | *Treponema pallidum*

1 pyruvate
2 citric acid
3 oxaloacetate

Metabolic modules with special importance in biotechnology



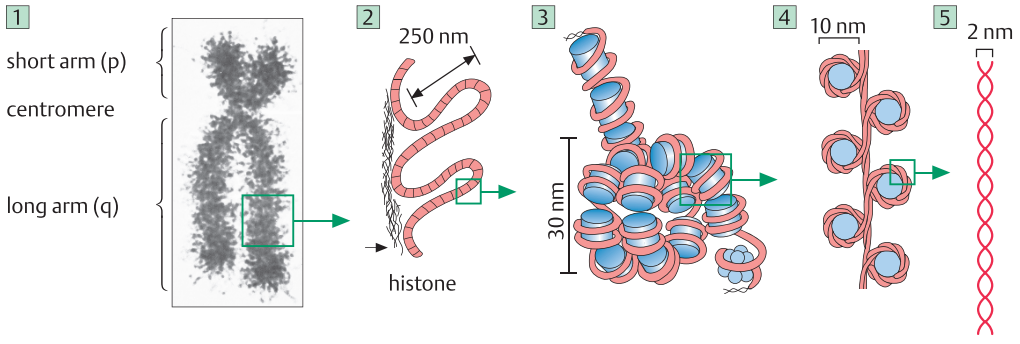
DNA: structure

General. During cell division, the genetic information of a cell is transferred from a parental cell to daughter cells (in prokaryotic cells: during cell division; in eukaryotic cells: after fusion of two haploid parent cells). The chemical substance containing this information is deoxyribonucleic acid (DNA), a supramolecular double helix of M_W up to 10^9 Da, composed of two single molecules. Most living beings store their genetic information in DNA molecules. This makes its heterologous transfer among different, unrelated species in principle possible, although such events are rare, e. g., during viral infection. Since the early 1970s, technical methods have been developed that allow the transfer of genetic information among different organisms (genetic engineering). These new techniques have led to a revolution in cell biology and to major advances in biotechnology.

Chemical structure of DNA. The individual building blocks of DNA are termed nucleotides. Their structure has two components: a deoxyribose-5'-phosphate moiety and one of the 4 bases adenine (A), guanine (G), thymine (T) or cytosine (C), which are glycosidically linked via their N1 nitrogen to position 1' of the sugar moiety. In DNA, nucleotides are linked together as a sugar phosphodiester polymer by a phosphate bridge between the 5'-C atom of one and the 3'-C atom of a second nucleotide. Such polymers can hybridize in a highly specific manner to a supramolecular double helix, if the base sequence allows for the sequential formation of either 2 hydrogen-bonded AT or 3 hydrogen-bonded GC base pairs. As a consequence, only two single-strand polymers of DNA, which are (largely) complementary in their nucleotide sequence, may form a double helix. DNA isolated from organisms has the following properties: 1) its molar mass is extremely high, 2) the genetic information is stored in its linear sequence of nucleotides, 3) the two single-strand polymers are unidirectional, i. e., they each contain one 3'- and one 5'-terminus, and 4) both strands may serve as a template for transferring its sequence information to a copy with a complementary nucleotide sequence.

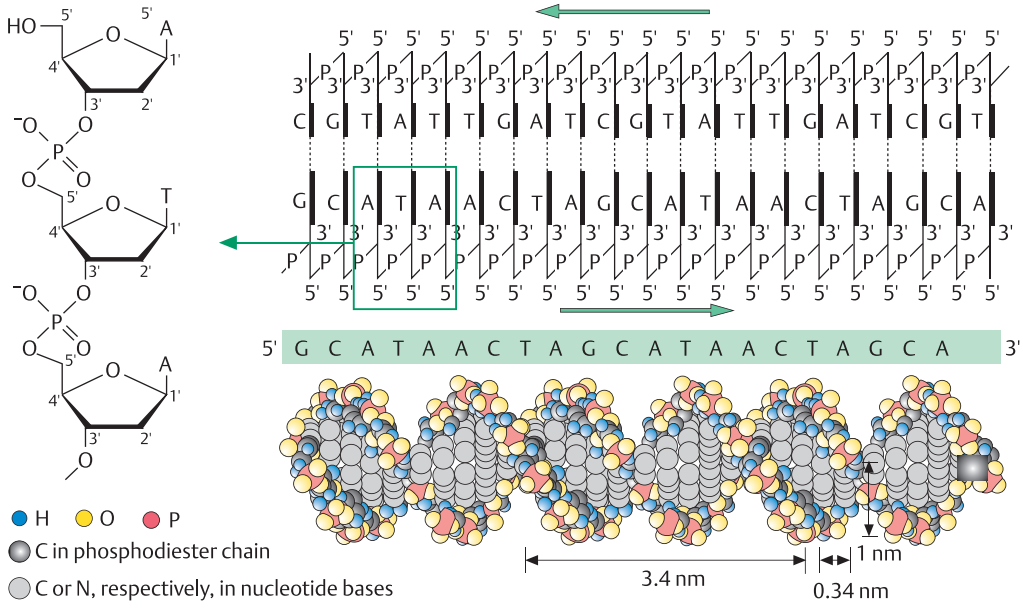
Structure of DNA in organisms. The total DNA of an organism is termed its genome. The large size of the DNA molecule and its important function in the storage and inheritance of genetic information requires special subcellular structures. In higher organisms DNA is usually distributed among several chromosomes whose number does not depend on genome size. Thus, bakers' yeast (*Saccharomyces cerevisiae*) contains 16, the fruit fly (*Drosophila melanogaster*) has 4 chromosomes, and the much larger quantity of DNA in humans is distributed among 23 chromosomes. The chromosomes are localized in the cell nucleus, where they form chromatin [a complex of DNA with basic proteins (histones)]. The length of DNA is usually given as the number of base pairs (bp). Human chromosome 3, for example, contains 2 DNA strands of ca. 160 million bases each. Since 3×10^6 bp have a calculated length of 1 mm, the extended length of the DNA double helix of chromosome 3 would be ca. 5 cm. The human DNA of all 23 chromosomes in a single haploid egg or sperm cell has a combined mass of ca. 3×10^9 bp, corresponding to a molar mass of $M_W = 1.8 \times 10^{12}$ Da and a calculated length of ca. 1m. The ca. 10^{13} standard diploid cells of our body contain a double set of 46 chromosomes, equivalent to ca. 2m of DNA per cell. During each cell division, all 46 double strands are replicated and packaged again into 46 chromosomes in the daughter cells. The DNA of prokaryotes exhibits simpler packaging. Because prokaryotes do not contain a cell nucleus, their DNA molecule is contained in a subcellular region of the cytoplasm, the nucleoid. Even the genome of *Escherichia coli*, a single circular DNA double helix of M_W 2.4×10^9 Da, would have an extended length of 1.5 mm. The replication of this huge molecule in *E. coli* under favorable growth conditions does not take more than 20 min, the doubling time of this organism. It proceeds with extremely high fidelity (error rate ca. 10^{-7} per gene of 1000 bp). Some organisms are polyploid, i. e. they contain more than a double set of chromosomes per cell. Thus, trout are tetraploid, wheat is hexaploid, and strawberries are decaploid. Polyploidy originates from whole genome duplication and is frequent in cultivated vegetables, fruits or flowers such as orchids due to continuous inbreeding. Polyploids are usually unable to interbreed with their diploid ancestors.

From a chromosome to the double helix

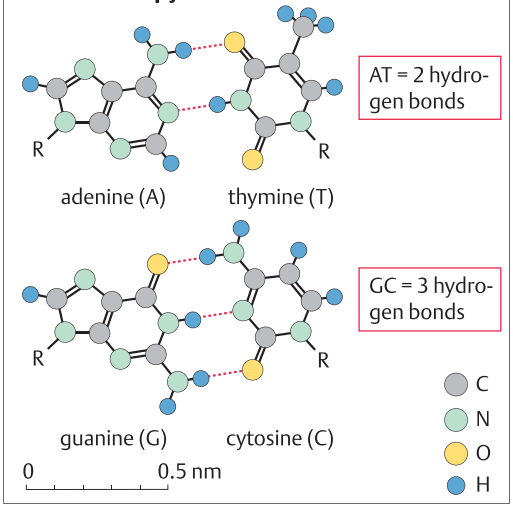


1 chromatin-histone loop
2 chromatin loops: ~100 000 bp (per loop)
3 30 nm fiber: ~1200 bp
4 nucleosome bead: ~80 bp
5 DNA double helix: 10 bp per 3.4 nm

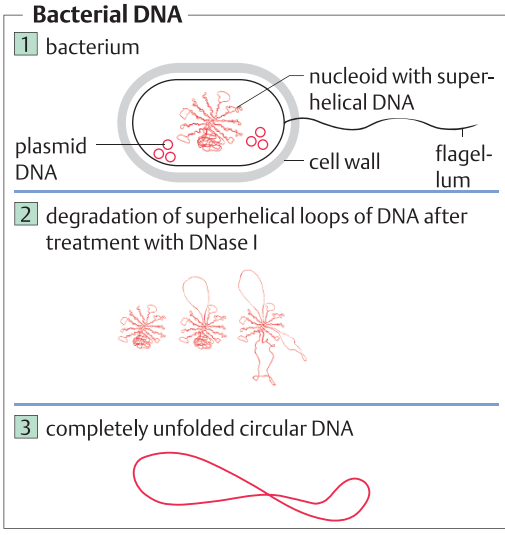
Structure of DNA



Purine and pyrimidine bases



Bacterial DNA



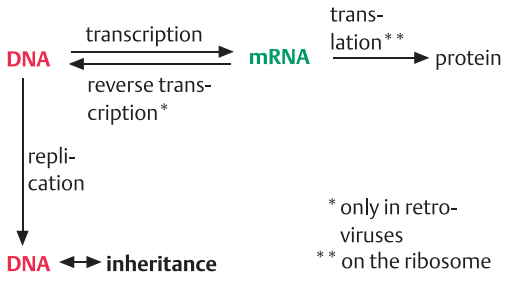
DNA: function

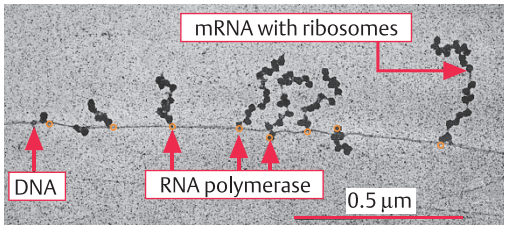
General. The information stored in DNA codes for the biosynthesis of proteins. In prokaryotic organisms, protein biosynthesis occurs in two steps: transcription of a DNA segment (usually a gene) into messenger RNA (mRNA), a polyribonucleic acid, and translation of mRNA into a protein sequence, using the machinery of the ribosomes. In eukaryotic organisms, this process is more complex. First, only part of the DNA, the exons, codes for the synthesis of proteins. In higher organisms, the noncoding sequences of the DNA (introns) may exceed exon DNA by a factor of 10 or more in length; the function of introns is still unclear. Within the cell nucleus, the DNA is transcribed into mRNA (primary transcript), followed by removal of the intron part of the transcript by splicing. The spliced, mature mRNA now leaves the nucleus and attaches to the ribosomes, which are located in the cytoplasm and on the endoplasmic reticulum. At the ribosomes, the protein chains are assembled according to the sequence information encoded in the mRNA. In some cases, the mRNA also encodes leader sequences, which direct a protein into a special compartment of the cell. Protein properties can be further modified through differential splicing during translation, removing one or several exons, or through posttranslational modification (e. g., glycosylation or phosphorylation). Such modifications in higher organisms may lead to a roughly 50-fold greater diversity of proteins than the number of coding genes; in humans, 20,300 genes may code for ca. 1 million protein variants, which differ among the types and age of cells. Such variations are analyzed by proteomics techniques (→314).

The genetic code used by living organisms to translate DNA sequences into protein structures is nearly universal. Each triplet of nucleotides of coding DNA, through the transcribed mRNA sequence, leads to highly selective incorporation of one distinct amino acid into the growing peptide chain. Due to the universal character of the genetic code, it is in principle possible to transfer genetic information from one organism to another. The genetic transformation of a host organism with DNA of foreign origin is the basis of genetic engineering technology (→44). In gene-transfer experiments, donor and host organism may show different preferences for some triplet codons. Thus, L-seleno-cysteine, a rare proteinogenic amino acid which occurs in some dozen pro- and eukaryotic proteins, is coded by UGA in a stem-loop mRNA structures. More importantly, the genetic code is degenerate: there are more unique triplet codes (A, T, G and C: 4^3 = 64) than amino acids (20), leading to a redundancy of up to 6 triplet sequences all coding for a single amino acid (e. g., UUA, UUG, CUU, CUC, CUA, and CUG all code for leucine). As shown in this example, the third base of the triplet contributes least to codon specificity (wobble hypothesis). However, the actual usage of a code type (in more precise words: the amount of each triplet-specific tRNA) differs among organisms and led to problems in early gene transfer experiments. Due to the rapid advances in genome sequencing (→56, 312), however, the codon usage of many organisms is now known, and it has become inexpensive and fast to have synthetic genes (→54) prepared (in 2014, for about 0.3 €/bp). As a result, genes with codons optimized for the host organism of choice or even whole gene clusters or gene modules required for metabolic engineering studies, are chemically synthesized.

Synthetic oligonucleotides (→54) are required for many experiments in genetic engineering. They are used as probes for hybridization experiments and as primers (starter sequences for DNA polymerase) in polymerase chain reaction (PCR) experiments and for site-directed mutagenesis of proteins. A well-selected probe or primer hybridizes with a highly specific sequence of the DNA isolated from an organism or a cell and allows for the detection, cloning, and amplification of an individual gene sequence. If no DNA sequence of the protein to be cloned is available, a putative sequence may be derived by sequencing the purified protein. In such cases, the fact that the genetic code is degenerate necessitates the chemical synthesis of many putative primers ("degenerate probe/primer") for the cloning experiment. If the probe or primer mixture leads in fact to hybridization or amplification (preferred method of identification: Southern blot (→60) or gel electrophoresis after amplification (→58)), the desired DNA can be identified in consecutive experiments, isolated and sequenced.

Functions of DNA





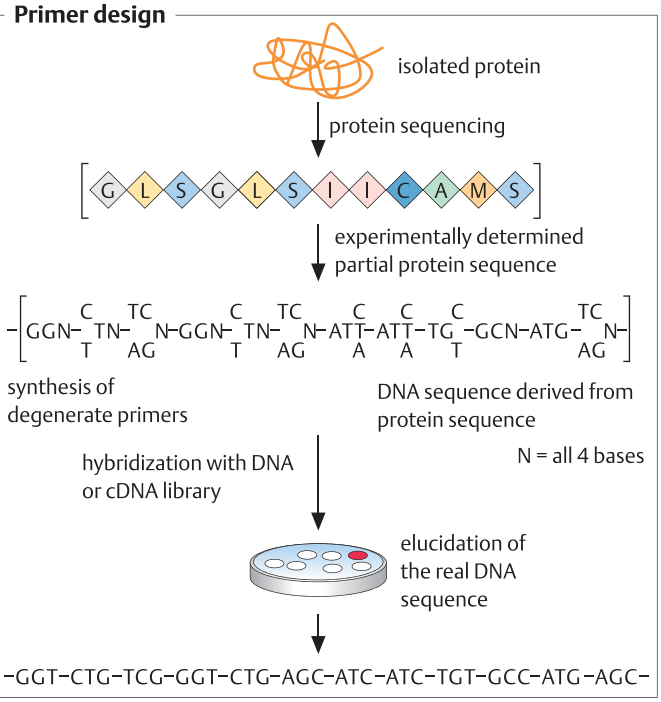
transcription and translation in a prokaryotic cell

The genetic code

amino acid	codon	amino acid	codon	amino acid	codon	amino acid	codon
phenylalanine (F)	UUU	serine (S)	UCU	tyrosine (Y)	UAU	cysteine (C)	UGU
	UUC		UCC		UAC		UGC
leucine (L)	UUA		UCA	**stop**	**UAA**	**stop**	**UGA***
	UUG		UCG	**stop**	**UAG**	tryptophan (W)	UGG
leucine (L)	CUU	proline (P)	CCU	histidine (H)	CAU	arginine (R)	CGU
	CUC		CCC		CAC		CGC
	CUA		CCA	glutamine (Q)	CAA		CGA
	CUG		CCG		CAG		CGG
isoleucine (I)	AUU	threonine (T)	ACU	asparagine (N)	AAU	serine (S)	AGU
	AUC		ACC		AAC		AGC
	AUA		ACA	lysine (K)	AAA	arginine (R)	AGA
methionine (M)	AUG		ACG		AAG		AGG
valine (V)	GUU	alanine (A)	GCU	aspartic acid (D)	GAU	glycine (G)	GGU
	GUC		GCC		GAC		GGC
	GUA		GCA	glutamic acid (E)	GAA		GGA
	GUG		GCG		GAG		GGG

*UGA in stem-loop structure: selenocysteine

Primer design



Codon usage

amino acid	co-don	frequency		
Glu, E	CAG	0.30	0.31	0.59
	CAA	0.70	0.69	0.41
Lys, K	AAG	0.24	0.43	0.60
	AAA	0.76	0.57	0.40
Pro, P	CCG	0.55	0.12	0.11
	CCA	0.20	0.42	0.27
	CCU	0.15	0.31	0.29
	CCC	0.10	0.15	0.33

E. coli
S. cerevisiae
man

RNA

General. It is widely held that the DNA-based genome, the genetic program of today's biosphere, was preceded by a simpler life form whose replication was based on RNA. The machinery used by cells for protein synthesis is largely based on ribosomal RNA, transfer RNA and messenger RNA. In biotechnology, RNA-based techniques play a considerable role. Examples are 1) RNA-based aptamers as bioaffinity molecules, 2) mRNA-based procedures to prepare proteins *in vitro*, 3) the capacity of interfering RNA to knock out gene functions, and 4) RNA-based vectors in gene therapy.

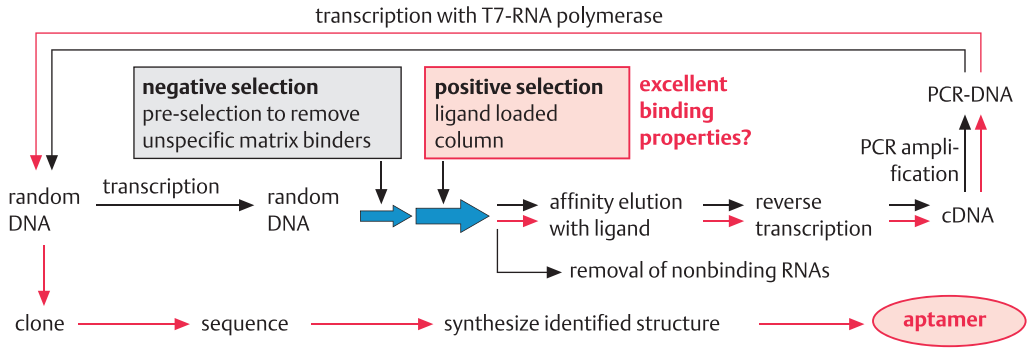
Aptamers are artificial nucleic acid ligands that bind with high affinity to hydrophilic molecules such as peptides and drugs. In the SELEX process (**s**ystematic **e**volution of **l**igands by **ex**ponential enrichment), vast combinatorial libraries of synthetic nucleic acids (10^{14}–10^{15} different molecules) are screened for binding to a target molecule. Those sequences that interact with the target are amplified by RT-PCR and transcribed in vitro, providing a sub-library after each round with further enhanced binding properties. Aptamers with binding capacities in the lower nM range have been isolated and have been studied both for diagnostic (→258) and therapeutic applications, for example, for selective gene inactivation. Aptamer arrays are also used in proteome analysis (SomaScan™).

Cell-free protein expression. Techniques have been developed for protein synthesis *in vitro*, starting from a DNA template. Using an optimized *E. coli* lysate that contains an RNA polymerase, tRNAs, ribosomes, amino acids, and ATP, up to a few mg of protein can be synthesized within 24 h. The method has been advantageously used to explore bottlenecks in transcription and to express proteins such as proteases or antibacterial peptides that are toxic to a host organism. Equipment for the use of this technology is commercially available (ProteoMaster™).

Knock out of gene functions. (→64) RNA molecules are involved in crucial steps of genetic information processing such as the splicing of exons and the synthesis of proteins. RNA interference, also termed post-translational gene silencing, has been recognized as a mechanism for regulating gene expression and mediating resistance to endogenous and exogenous pathogenic RNAs such as RNA viruses ("immune system of the genome"). In some of the above mechanisms, RNA can also be catalytically active (ribozymes), for example, splitting phosphodiester bonds in the absence of any protein. Many of these mechanisms are being explored for use in biotechnology. Antisense RNA – a technology that is discussed under the topic of gene silencing (→64) – has been successfully applied to eliminate polygalacturonase activity in ripening tomatoes (FlavrSavr®), thus leading to better aroma without wrinkling of the skin. *Trans*-cleaving ribozymes (i. e., ribozymes that cleave a foreign strand of RNA) have been explored in the therapy of HIV and breast and colorectal cancer up to the clinical phase II level. They can be applied, for example, by infusing transformed $CD4^+$ lymphocytes or $CD34^+$ hematopoietic precursor cells from the infected patient, which have been expanded *ex vivo*. In many eukaryotic cells, mRNA can be destroyed by a process termed RNA interference (RNAi): the presence of double-stranded RNA activates an RNase able to recognize and digest matching endogenous mRNA, possibly by using the double-stranded RNA as a template. It is based on the random cutting of the double-stranded RNA by an RNase (DICER); after ATP-dependent enzymatic activation of the single-stranded fragments generated in this process, they can hybridize specifically with the mRNA and be recognized by an RNase, which then degrades the mRNA. For example, when a suitable synthetic double-stranded RNA is expressed behind an RNA polymerase III (Pol III) promoter, HIV gene expression in cotransfected cells is largely inhibited. Interfering RNA is now widely investigated for drug development. A major issue here is how to deliver an antisense molecule to its target, and how to counter inactivation by interferons.

Gene therapy. The use of viral RNA vectors for human gene therapy is described elsewhere (→304). mRNA extracts of human tumors were successfully used to transform monocytes of the same patient, resulting in mature dendritic cells loaded with tumor-specific RNA which, upon infusion, stimulated the immune system of the patient to form anti-tumor cytotoxic T-lymphocytes.

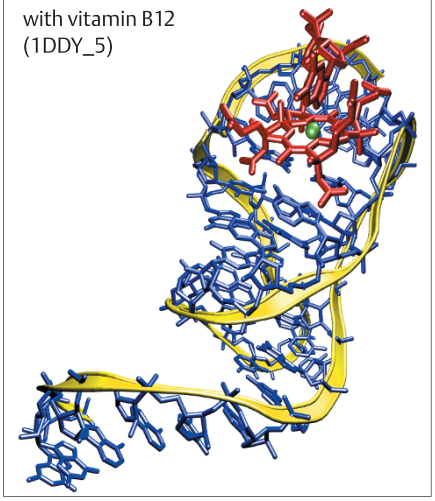
Steps in aptamer selection (SELEX process)



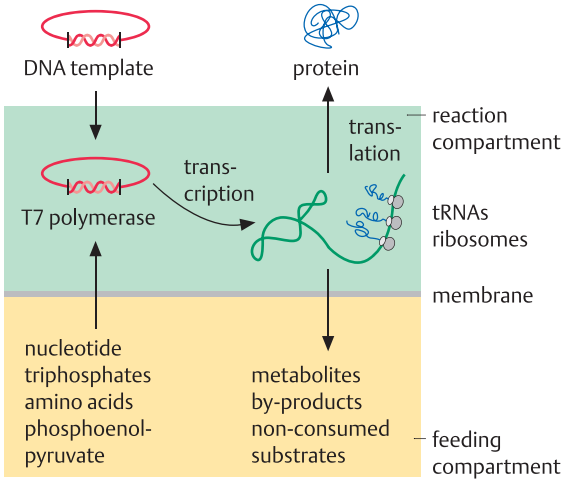
Some RNA-derived aptamers binding to extracellular proteins

target	Kd(nM)	nucleic acid
acetylcholine receptor	2.0	RNA
basic fibroblast growth factor (PDGF)	0.35	RNA
interferon-γ	6.8	2'-modified RNA
keratinocyte growth factor (KGF)	0.0003	2'-modified RNA

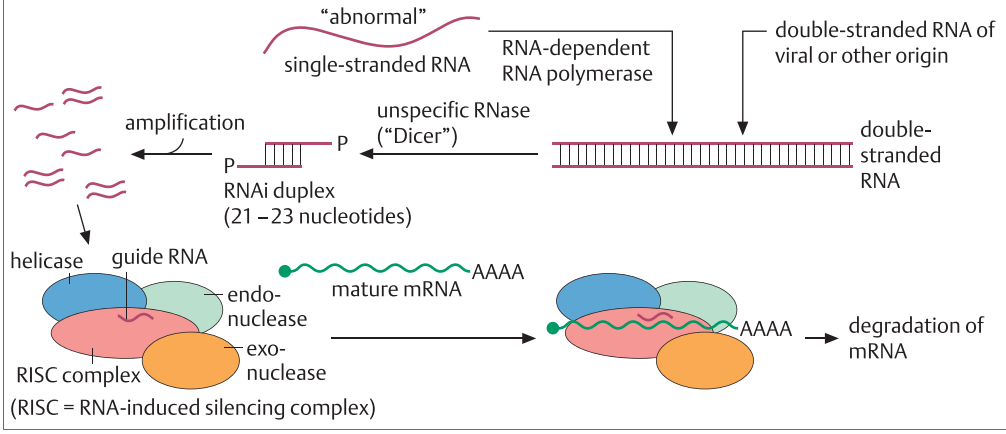
Aptamer complex



Biosynthesis of proteins *in-vitro*



RNAi – interfering RNA



Genetic engineering: general steps

General. Although genetic engineering has a wide range of applications, only a few fundamental steps are required to transfer and express foreign DNA in a host cell. These include 1) manipulation of DNA, in particular its isolation, amplification, enzymatic modification, characterization, sequencing, and chemical synthesis, and 2) cloning and expression of the DNA in pro- and eukaryotic cells.

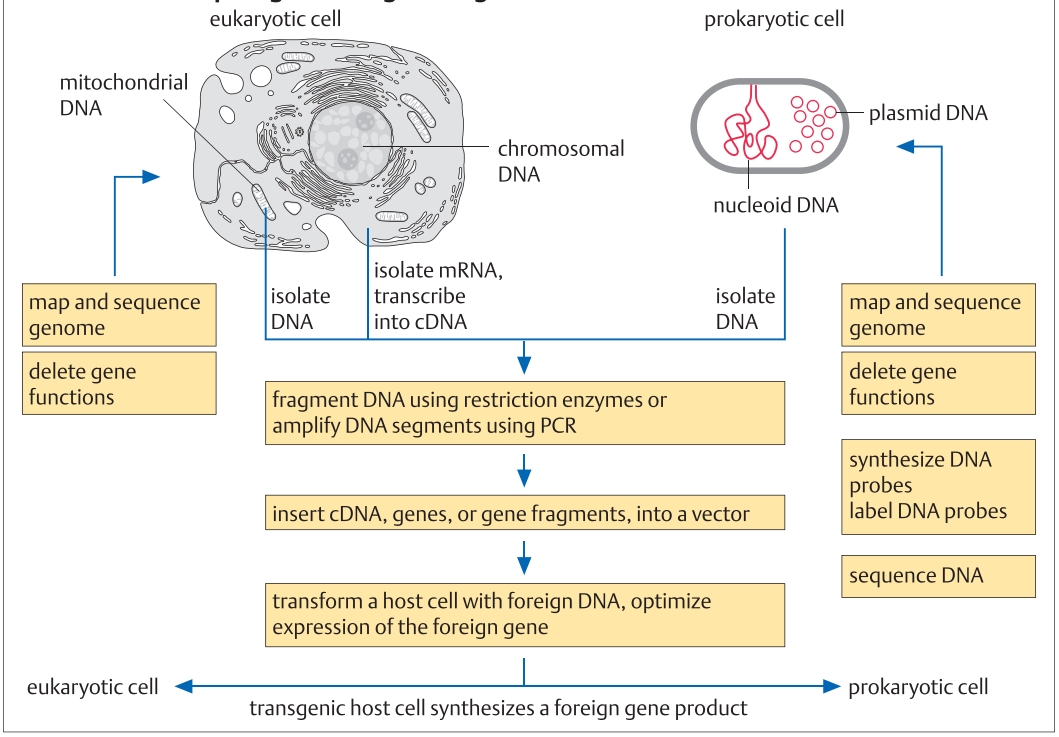
Purification, enzymatic modification, and amplification. DNA is found in each cell in very small quantities (just one molecule in a prokaryotic cell and in each chromosome of a eukaryotic cell), but it can still be isolated in pure form by extraction. For the next steps required in genetic engineering, intact DNA is too large. Fortunately, enzymes have been found that selectively cut, modify, ligate, or amplify DNA. Other enzymes help to transcribe DNA sequences into mRNA in a test tube ("*in vitro transcription*"). DNA fragments can be chemically synthesized, using a robotic device. For most tasks, the amount of DNA fragments that can be obtained from a cell is not enough. It is therefore very important that DNA segments up to a length of 1000 bp or more can be amplified using the polymerase chain reaction (PCR) (→50). This procedure allows, for instance, recombining DNA fragments of different origins (e. g., from different organisms or from chemical synthesis).

Characterization and sequencing. A DNA fragment can be characterized according to 1) its melting behavior (a DNA double helix containing a high amount of GC base pairs, with 3 H-bonds, melts at a higher temperature than AT-rich DNA, with 2 H-bonds), 2) its molar mass, which is usually determined by gel electrophoresis, 3) its nucleotide sequence, 4) biological characterization of its functional elements, and 5) mapping of those sequences that can be enzymatically cut by diverse but selective restriction enzymes.

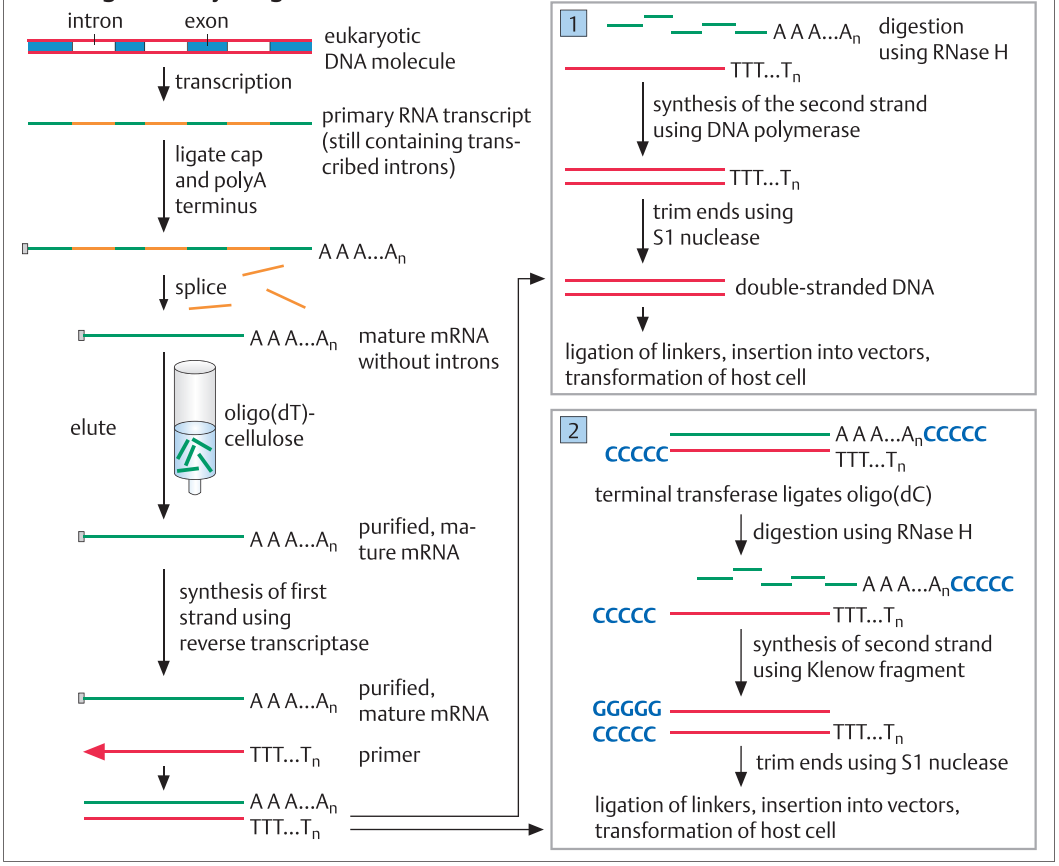
Cloning. A DNA sequence contained in the genome of a prokaryotic organism, usually a gene that codes for a protein, can be directly cloned. If a functional gene from a eukaryotic cell is to be cloned, the introns must first be removed from the DNA. This need is usually circumvented by starting with mature mRNA, i. e., mRNA after splicing, and transcribing it in vitro into complementary double-stranded DNA (cDNA), using the enzyme reverse transcriptase (RT), which is mostly isolated from mammalian retroviruses. Heat-stable DNA polymerase from the thermophile *Thermus thermophilus* can also be used in this reaction (→196). In the presence of Mn^{2+}, this enzyme functions like a reverse transcriptase and accepts RNA as well as DNA as a template. The mRNA-DNA hybrid that results from this reaction is then degraded to single-stranded DNA using RNase and serves as a template in a standard PCR reaction to generate double-stranded DNA. Another feature of eukaryotic mRNA isolated from a cell can be exploited to amplify it in an unspecific manner: because all eukaryotic mRNA molecules carry 3' sequences of polyadenine (polyA), a complementary polythymidine oligomer (polyT) can be used as a primer for enzymatic synthesis with RT.

Expression. Several methods are available, and are described later, for cloning and replicating DNA, and expressing (transcribing and translating) foreign DNA or cDNA in a host cell (→58). Several types of bacteria are favored as host cells for the following reasons: 1) their genome consists of a single DNA double helix, 2) their molecular genetics have been thoroughly studied, 3) phages or plasmids are available that can be used as vectors for the introduction of foreign DNA, 4) they propagate rapidly and can be bred in large quantities in a bioreactor, 5) they are non-pathogenic or can be easily and irreversibly transformed into non-pathogenic mutants. The preferred species for the cloning and expression of foreign DNA are attenuated strains of *Escherichia coli*, e. g., *E. coli* K12 (→20). However, other bacteria, such as *Bacillus subtilis*, Lactobacilli, and various Pseudomonas and Streptomyces strains, have also been successfully used. Higher organisms such as fungi (Aspergillus, Trichoderma) (→16), yeasts (Saccharomyces, Schizosaccharomyces) (→14), mammalian cells (CHO cells, BHK cells) (→98), insect cells (transformed by baculovirus), whole animals (such as silkworms for protein production or transgenic mice as human disease models) (→272), or transgenic plants (for the production of materials such as oils, starch or biopolymers) (→284), have also been transformed by foreign DNA.

Fundamental steps in genetic engineering



Cloning of eukaryotic genes



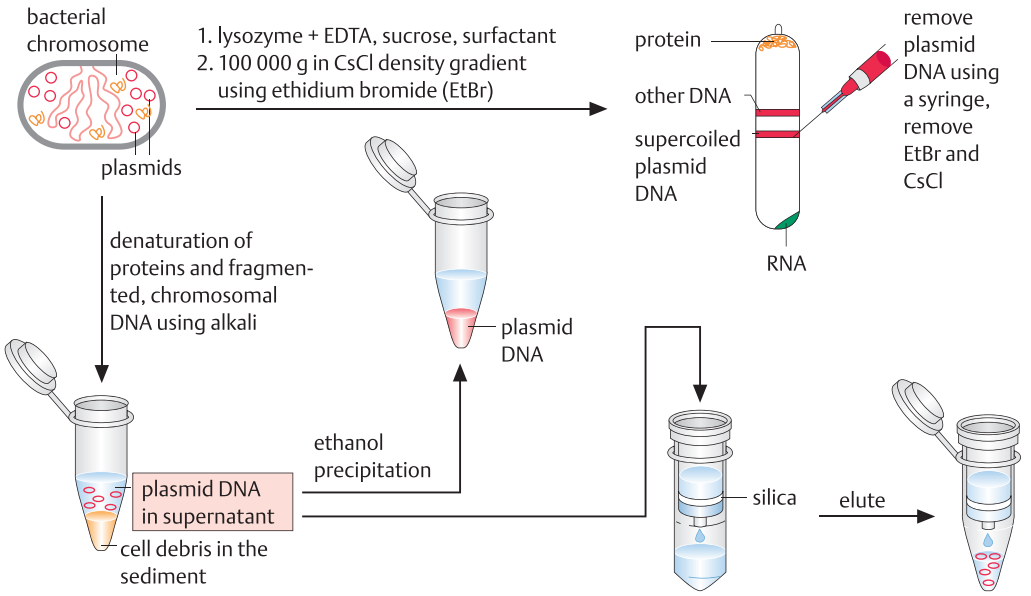
Preparation of DNA

General. The preparation, modification, and characterization of DNA from living cells are usually the first steps in genetic engineering. Restriction endonucleases are an essential tool in such experiments. They are also used for constructing physical gene maps.

Preparation of DNA. Depending on the organism or cell type, DNA may occur in different forms, requiring various protocols for its isolation. Prokaryotic DNA is not contained in a cell nucleus. It can be isolated after lysis of the cell wall with a lysozyme/detergent mixture (usually sodium dodecylsulfate, SDS), and denaturation of the cell proteins by phenol/chloroform. If this mixture is centrifuged, DNA can be precipitated from the supernatant by the addition of cold ethanol. Frequently, bacteria contain not only a single molecule of genomic DNA, but also plasmid DNA, which is of much lower molecular mass and is very important in genetic experimentation. If cell lysis is followed by the addition of NaOH and detergent, proteins are precipitated and chromosomal DNA is partially hydrolyzed. Under suitable conditions, plasmid DNA is preserved as a circular double strand and can be isolated, from the supernatant of a centrifugate, by centrifugation in a density gradient (sucrose or CsCl), by ethanol precipitation, or, more simply, by chromatography. The method of choice is often solid-phase chromatography on mini columns (spin columns). In this procedure, DNA is first bound to silica in the presence of high salt concentrations and eluted at low salt concentrations. This fast and inexpensive method delivers high-purity DNA ready for use in downstream applications. It can easily be automated by using magnetized silica particles,. Depending on the amount of isolated DNA, this last procedure is termed miniprep (ca. 20μg DNA) or maxiprep (several mg). Phage DNA is enriched from infected bacterial cultures by removing the lysed bacteria by centrifugation, precipitating the phage particles with polyethylene glycol, removing the phage capsid by phenol extraction, and precipitating phage DNA from the supernatant with ethanol. Eukaryotic DNA is distributed among several chromosomes that are contained in the cell nucleus. Total DNA from animal cells is obtained by lysing the cells with detergent, digesting proteins and RNA with proteinase K and RNase, removing detergent and proteins by salt precipitation, and, finally, precipitating DNA from the supernatant with ethanol. In studying eukaryotes, spliced mRNA is used as a source of genetic information. mRNA is isolated from organs or cell cultures by lysing the cells, removing the cell nuclei by centrifugation, removing cytoplasmic proteins by phenol extraction, and precipitating the mRNA from the supernatant. Mitochondria and the plastids of eukaryotic photosynthetic organisms contain their own DNA, which replicates independently of the chromosome. It can be isolated by analogous procedures from these organelles after their isolation by differential centrifugation. Special types of DNA extraction procedures are required for "ancient DNA" that is partially degraded, and for samples containing inhibitors that interfere with subsequent analytical procedures such as humic acids or hemoglobin in forensic DNA analysis.

Restriction enzymes (restriction endonucleases) are synthesized by bacteria as protection against foreign DNA (e. g., phage DNA) that may enter their cytoplasm. They are widely used in genetic engineering to selectively cut the long native DNA strands into smaller, well defined fragments. They bind to a recognition sequence of DNA that is 4–10 nucleotides long and then hydrolyze the DNA double strand within or near the recognition site so as to leave blunt or partially single-stranded (sticky or cohesive) ends. If a purified DNA molecule is hydrolyzed by different types of restriction enzymes and the size of the fragments from each experiment is recorded by gel electrophoresis or DNA sequencing, a restriction map of the DNA is obtained, which can be used as a strategic basis for inserting and expressing foreign DNA. The frequency of cleavage sites for a given restriction enzyme can be statistically calculated from the length of the recognition sequence; thus, an enzyme with a 4-bp recognition site (e. g., *Alu* I, AGCT) will make cuts at $(¼)^4 = (1/256)$, or once every 256 bp. However, the sequence information in DNA is not randomly distributed, so the observed cleavage frequency differs from this calculation. Recognition also depends on the G+C content of a DNA and thus can be used for characterizing an organism.

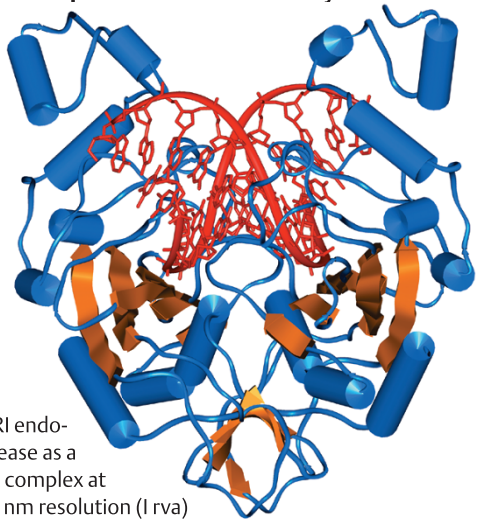
Isolation of bacterial DNA



Restriction enzymes (endonucleases)

enzyme	organism	recognition sequence (5' →3')	3'-end
Not I	***N****ocardia* ***ot****itidis-caviarum*	GCGGCCGC	sticky
Eco RI	***E****scherichia* ***co****li*	GAATTC	sticky
Bam HI	***B****acillus* ***am****yloliquefaciens*	GGATCC	sticky
Bgl II	***B****acillus* ***gl****obigii*	AGATCT	sticky
Pvu I	***P****roteus* ***vu****lgaris*	CGATCG	sticky
Pvu II	***P****roteus* ***vu****lgaris*	CAGCTG	blunt
Hind III	***H****aemophilus* ***in****fluenzae* R_d	AAGCTT	sticky
Hinf I	***H****aemophilus* ***in****fluenzae* R_f	GANTC	sticky
Sau 3A	***S****taphylococcus* ***au****reus*	GATC	sticky
Alu I	***A****rthrobacter* ***lu****teus*	AGCT	blunt
Taq I	***T****hermus* ***aq****uaticus*	TCGA	sticky

DNA complex of a restriction enzyme



Eco RI endonuclease as a DNA complex at 0.20 nm resolution (I rva)

Cleavage frequency

cleavage of λ-DNA using	number of cleavage sites calculated	found
Bgl II	12	6
Bam HI	12	5
Sal I	12	2

all 3 restriction enzymes recognize sequences of 6 base pairs. This leads to a statistical cleavage frequency of one every 4^6 or 4 096 base pairs, corresponding to 12 cleavage sites for the ~49 000 bp of λ phage. However, significantly fewer cleavage sites are observedexperimentally, because the sequences required for hydrolysis are not randomly distributed over the phage genome

Other useful enzymes for DNA manipulation

General. For manipulation of DNA in vitro, the following enzymes have found wide use: 1) hydrolases that cut DNA at specific sequences, 2) lyases, which ligate DNA fragments to each other, 3) synthetases that polymerize a DNA double strand on a single-strand DNA template, 4) hydrolases and transferases that modify DNA at the 3' or 5' end, and 5) methyl transferases which protect synthetic DNA fragments from attack by bacterial restriction endonucleases and are important enzymes in epigenetic processes.

Hydrolases. The most important hydrolases are the restriction endonucleases, which were discussed above. A wide choice of enzymes is available commercially, differing in their recognition sequences and in their ability to generate blunt or sticky ends (i. e., ends with a single-strand overhang a few nucleotides long). Another frequently used hydrolase is nuclease S1, obtained from *Aspergillus niger*: it cleaves single-stranded DNA and double-stranded DNA at single-stranded gaps.

Lyases, transferases. The most important enzyme of this group is DNA ligase. It functions in the cell as a repair enzyme: it repairs gaps that may have occurred in a double-stranded DNA molecule during replication, recombination, or accidental events. This important enzyme, which occurs in all cells, is used in genetic engineering to recombine DNA fragments in vitro (e. g., to insert foreign DNA into a DNA vector). DNA ligase can ligate both blunt and sticky ends. However, because the ligation of sticky ends is much more effective (because the single-stranded sequence facilitates hybridization with the complementary sequence to be ligated), blunt ends are often transformed into sticky ends before ligation. This can be done by using linkers or adapters or by adding polymer tails (tailing) in the presence of the enzyme terminal deoxynucleotidyl transferase. A detailed description of all individual steps used in such protocols can be found in the pertinent literature. DNA ligase is usually isolated from *E. coli* cells that were infected with bacteriophage T4 (T4 DNA ligase), terminal transferase is obtained from calf spleen.

Synthetases. The two most important synthetases are DNA polymerase I and reverse transcriptase. DNA polymerase I is obtained from *E. coli*. In the presence of a single-stranded DNA primer, it adheres in *E. coli* cells and in the test tube to single-stranded DNA and synthesizes the second strand in the 5'→3' direction. In addition to its polymerase activity, it also shows 3'→5' and 5'→3' exonuclease activity. It also recognizes short stretches of single-stranded DNA (gaps) in an otherwise double-stranded DNA molecule, using them as a template for polymerization. Removal of the first 323 amino acids of this enzyme results in loss of the 5'→3' nuclease activity, but the resulting enzyme fragment (Klenow fragment) still has polymerase and 3'→5' nuclease activity and can fill in single-stranded gaps in an otherwise double-stranded DNA. The Klenow fragment is often used to introduce radioactive labels into DNA, a useful technique for visualizing traces of DNA by autoradiography (by exposure to an x-ray film) (→60). In labeling experiments, the preferred method is to use sticky 5' ends of the DNA fragment as a template, complementing it with ^{32}P-labelled deoxynucleotides in the presence of Klenow fragment. DNA polymerases from thermophilic microorganisms are widely used in PCR reactions. Reverse transcriptase (RT) is a key enzyme in retroviruses. It uses viral RNA as a template in synthesizing complementary DNA in a host cell. This property can be used in genetic engineering to synthesize cDNA from mRNA. RT is isolated from animal cells that have been infected with retroviruses such as the monkey myoblastoma virus (MMV) or the Moloney mouse leukemia virus (MMLV).

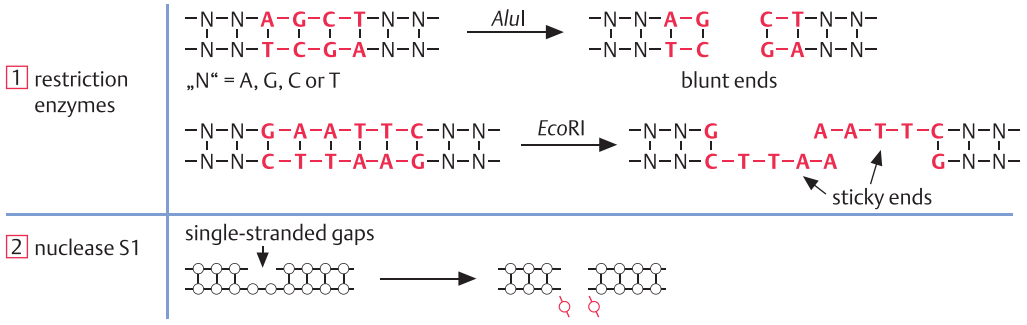
Endgroup-modifying enzymes. It is often necessary during cloning experiments to add or remove a terminal phosphate group at the 5' position of a DNA fragment. For this purpose, alkaline phosphatases and polynucleotide kinase are commercially available.

Methyl transferases modify DNA by covalent binding of methyl groups to cytosine and adenine (see Epigenetics) (→66). Products are 5-methylcytosine, *N6*-methyladenine und *N4*-methylcytosine. Synthetic DNA fragments can be protected through methylation from attack of bacterial restriction endonucleases.

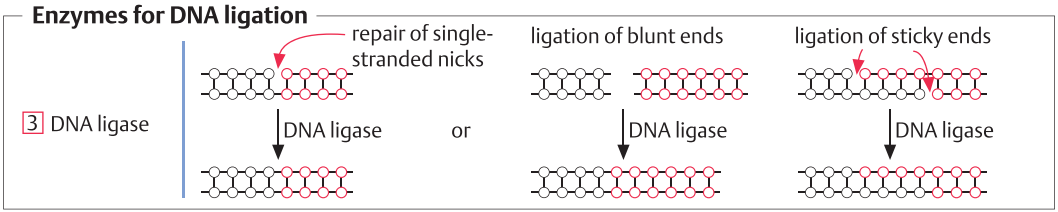
Enzymes for the manipulation of DNA

enzymes used in	function
DNA degradation	
restriction enzyme, endonuclease, [1] nuclease S1 [2]	cleaves internal phosphodiester bonds cleaves single-stranded DNA or single-stranded gaps
synthesis or ligation of DNA	
DNA ligase [3]	repairs single-stranded breakages, ligates two DNA molecules
DNA polymerase I [4]	synthesizes double strand, fills gap in single strand
Klenow fragment [5]	fills gaps in single strand, no 5' – 3' exonuclease activity
reverse transcriptase [6]	synthesizes DNA on a RNA matrix
modification of end groups	
alkaline phosphatase	removes phosphate group from 5' end
polynucleotide kinase	adds phosphate group to 5' end
terminal deoxynucleotide transferase	adds phosphate group to 3' end

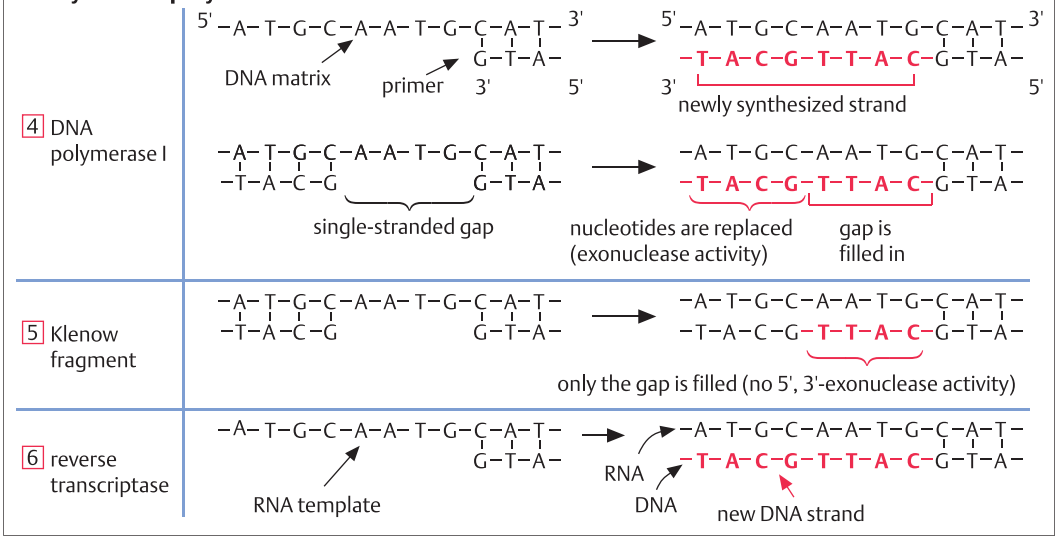
Enzymes for DNA cleavage



Enzymes for DNA ligation



Enzymes for polymerization of DNA on a matrix



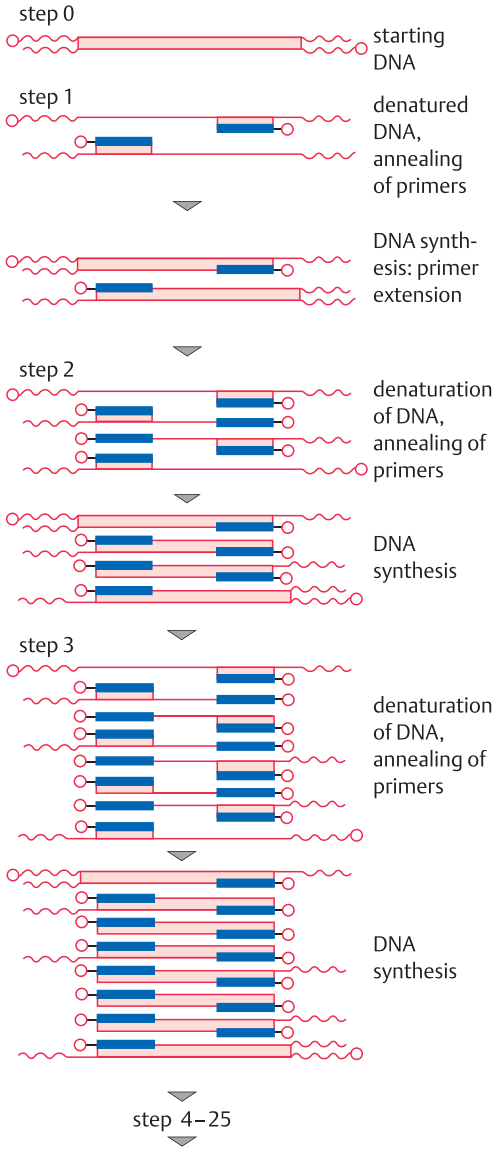
PCR: general method

General. The polymerase chain reaction (PCR) is one of the most important practical developments in genetic engineering, and its inventor, Kary Mullis, was honored with a Nobel Prize. Using PCR, any short sequence of DNA, e. g., a gene or gene fragment, can be amplified many times by using DNA polymerase. This technique is most valuable for the identification and manipulation of genes.

Method. The standard protocol requires two oligonucleotides (primers), which bind to either end of the DNA target sequence that is to be amplified (one for each DNA strand). This implies that these DNA sequences are either already known or that they can be determined from a protein sequence (if so, the ambiguities of the degenerate genetic code must be considered). Besides the DNA template and the two primers, a mixture of the 4 deoxynucleotides and a DNA polymerase is required. The PCR reaction then proceeds in three steps: 1) at 94°C, the DNA double strand is melted, forming two single-stranded DNA molecules (denaturation), 2) after lowering the temperature to 40-60°C, the two primers each hybridize to a DNA strand (annealing), and 3) after increasing the temperature to 72°C, two new complementary strands are formed (extension). If heating to 94°C is repeated, the newly formed DNA strands come off their DNA template, and the reaction cycle is repeated after cooling. With an automatic thermocycler, this cycle is repeated 25-40 times (a few seconds to several minutes per cycle, depending on the template), leading to amplification of the original DNA to between 2^{25} and 2^{40} copies within just a few hours. A condition for the reaction is that the DNA polymerase used must tolerate the high temperature required for melting the two DNA strands without becoming inactive. DNA polymerase from thermophilic microorganisms, e. g.. from *Thermus aquaticus, Pyrococcus furiosus*, or *Thermotoga maritima* (*Taq, Pfu*, or *Tma*-polymerase) have such properties. The error rate (mutation frequency per bp per amplification cycle) of *Taq* polymerase is about 8×10^{-6}. The two other polymerases are even more precise, because they also have proofreading activity. Thus, the error rate of *Pfu Ultra*, a *Pfu* polymerase optimized by protein design, is just 4.3×10^{-7} errors per base pair. The molar mass and yield of PCR products are determined by gel electrophoresis or, in real time, by including reporter groups such as SYBR Green (light cyclers) – a fluorescent dye that binds to double-stranded DNA. As the amount of DNA produced in a PCR reaction increases, the amount of fluorescence from the dye increases proportionally, and the original amount of DNA can be obtained by extrapolation.

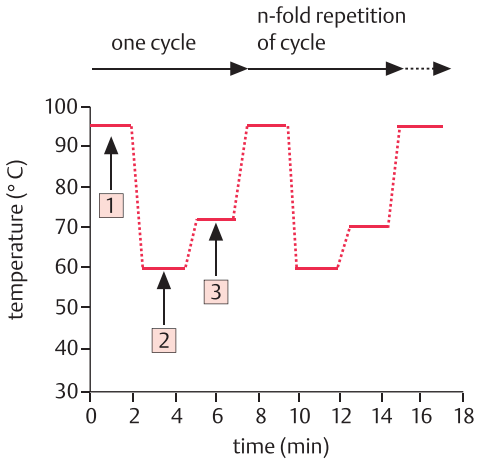
Practical applications. Using PCR, defined sequences of DNA can rapidly be cloned and sequenced. Because even single DNA molecules can be amplified by PCR, as has been shown by the amplification of DNA from a single sperm cell, the method has been applied in archeology and paleontology (→302). In clinical diagnostics, it can be advantageously used, once a correlation between a DNA sequence and a disease has been established (→300). This is already true for many infectious and, increasingly, also for genetically determined diseases. In the fields of food and environmental analysis, PCR methods help to identify traces of transgenic plant materials or of pathogens. Once the consensus sequences of a protein family are known, primers can be designed to help identify still unknown members of the family (reverse genetics). By using modified primers or deliberately increasing the error rate of a PCR reaction, defined or statistical mutations can be introduced into a protein. RNA is also amenable to PCR analysis after it has been transcribed into cDNA. It can be amplified (RT-PCR), and then used to, for example, determine 1) the load of RNA viruses in a cell (e. g., HIV virus), or 2) the relative quantities of mRNA in a cell. PCR techniques have been successfully miniaturized to proceed in micromachined capillaries ("PCR on a chip"). Using appropriate microdevices, a desired sequence of DNA may be amplified by a factor of 2^{20} in < 1h, to be further used for diagnostic assays, e. g., in a DNA array. Sequence-specific fluorophore-labeled probes allow for the simultaneous detection of several genes or gene fragments during "multiplex" assays, and automation in 96- or 384-well plates in combination with sequence-specific probes permits high-throughput multiplex analysis of genes in clinical samples, e. g. for the detection of single-nucleotide polymorphisms (SNPs). (→298, 300)

Polymerase chain reaction (PCR)



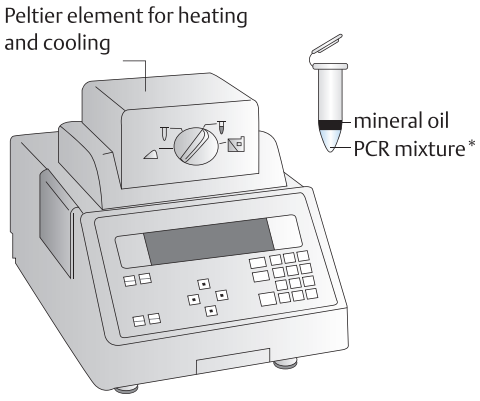
analysis of PCR products	
method	**analysis**
polyacrylamide gel electrophoresis	**ethidium bromide staining (UV light, image analysis)**
agarose gel electrophoresis	hybridization with labeled probe (Southern blot) incorporation of radioactive tracer, e. g., ^{32}P phosphate silver staining
HPLC	UV analysis
gel electrophoresis or HPLC after degradation with restriction enzymes	see above

Functional principle of a thermocycler

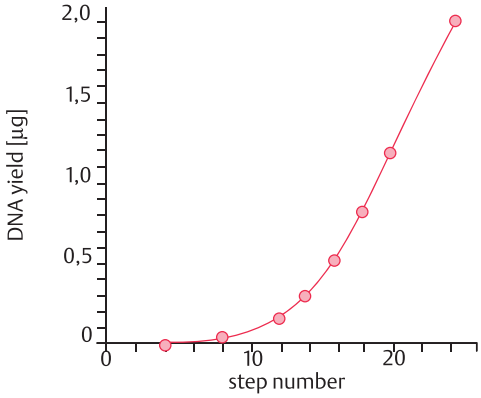


1 step 1: denaturation
2 step 2: annealing of primers
3 step 3: extension of primers

Thermocycler



* up to 96 parallel assays



by incorporation of ethidium bromide or SYBR Green, the PCR yield can be determined in real time (Light Cycler™)

PCR: laboratory methods

General. PCR is a key protocol for a wide range of molecular genetic experiments. Only a limited number of applications can be discussed here.

Incorporation of functional elements. Functional elements of DNA include cloning sites (recognition sequences for restriction enzymes), start or stop codons, and tags (sequences that serve to identify a protein after translation, or sequences that encode an N- or C-terminal polyhistidine, allowing for rapid purification of the translated protein by metal affinity chromatography) (→106).

Sequence analysis of mRNA (RT-PCR). mRNA can be amplified by PCR after its (partial) sequence is known directly or from the corresponding protein sequence. A pertinent primer is synthesized, annealed to mRNA isolated from a cell, and translated into the first single strand of cDNA using reverse transcriptase and a nucleotide mixture. In the next step, the single-stranded cDNA can be amplified by a standard PCR protocol, resulting in the amplification of specific sequences. RT-PCR is done in either one or in two steps including exchange of buffers.

Fusion of two DNA fragments. If two fragments of genes are to be fused, the desired sequences are amplified in two separate PCR steps, using a set of two primers. This results in PCR products that contain overlapping sequences at the desired fusion positions. In a third PCR reaction, the two PCR products are used as templates, adding the terminal primers, which leads to hybridization of the complementary strands with overlapping sequences and amplification of the fusion product. In this protocol it is important that the reading frame for the desired triplet is correctly chosen. It also may be necessary to insert a spacer between the two coding genes, e. g., a sequence coding for polyalanine. Such spacers may help preserve the free mobility of each of the two fused proteins (single-chain antibodies are a relevant example).

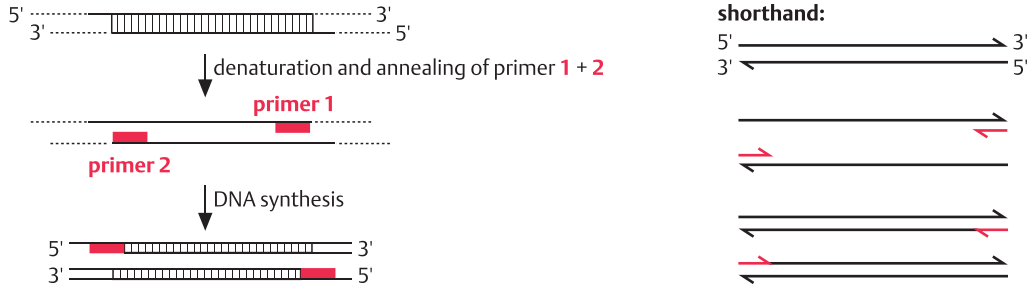
Insertion or removal of a gene segment. By analogy to gene fusion, a skillful choice of primers for internal or terminal sequences may lead to truncated DNA (and proteins), from which a desired segment has been deleted.

Site-directed mutagenesis (→198) is a very useful technique, for example, for elucidating enzyme mechanisms or for the targeted modification of an enzyme's substrate specificity. An older method for site-directed mutagenesis is based on introducing mutations into the single-stranded DNA of the M13 phage. It has been completely replaced by PCR protocols. Because DNA fragments can hybridize even when there is a mismatch between single nucleotides, a modified triplet code leading to the desired amino acid substitution can be introduced into any position of the DNA under study and can be amplified by PCR. Another method uses two complementary oligonucleotides, carrying the mutation, and a double-stranded plasmid, composed of permethylated DNA as a template. Using suitable primers and Pfu polymerase, the complete plasmid is amplified *in vitro*. The methylated template DNA (the DNA amplified *in vitro* is not methylated) is removed by digestion with the restriction endonuclease *Dpn*I, which hydrolyzes only methylated DNA. The newly synthesized DNA that carries the mutation can be directly transformed into *E. coli*, and time-consuming cloning steps are no longer necessary. Mutation kits based on this principle are commercially available.

Multiplex PCR. Several gene sequences can simultaneously be amplified in a single PCR, by combining suitable primer pairs. If the primers are labeled with different fluorescent markers, quantitative RT-PCR of several genes becomes possible. This procedure has become quite important in the clinical analysis of, e. g., viral infectious diseases, fast detection of antibiotic resistance, or analysis of single-nucleotide polymorphisms (SNPs).

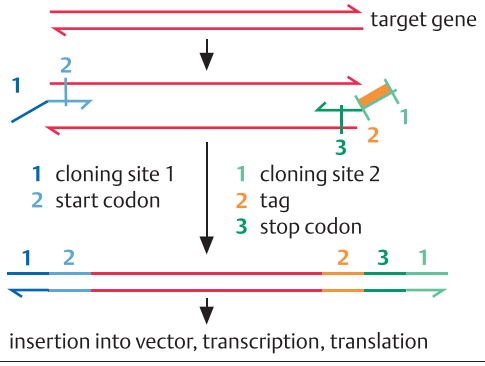
PCR with degenerate primers. Degenerate primers are families of homologous sequences of a single-strand DNA in which one or several of the nucleotides is substituted by an arbitrarily chosen nucleotide or by deoxy-inosine (dI). dI pairs to any of the other bases ("universal base") and thus can substitute for unknown nucleotides in a DNA sequence in a "degenerate primer." This allows genes whose precise sequence is unknown to be cloned, e. g., if the putative sequence was derived from a protein sequence and the codon usage is uncertain, or if genes of a multigene family are to be cloned.

Shorthand for PCR reactions

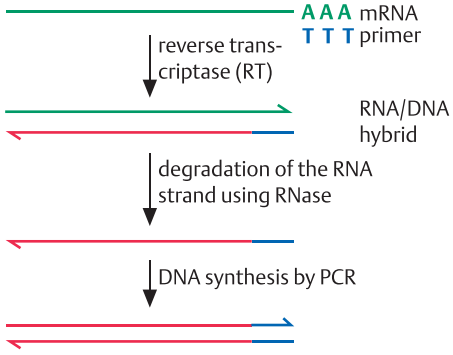


Application examples

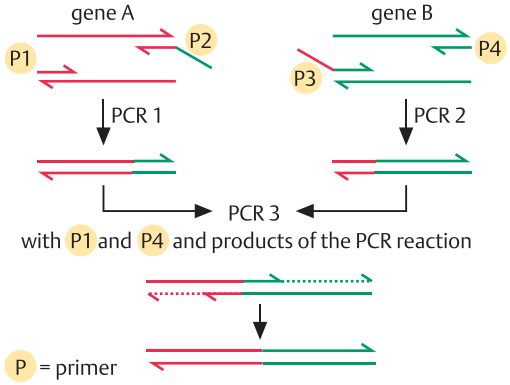
1 incorporation of functional elements



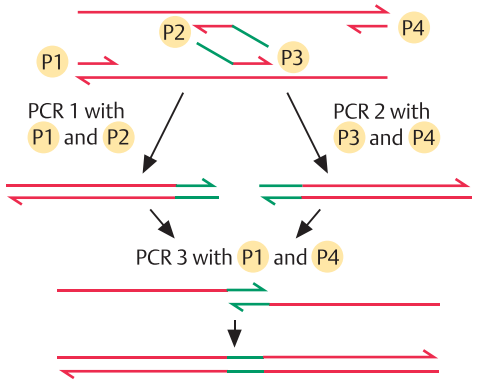
2 amplification of mRNA (RT-PCR)



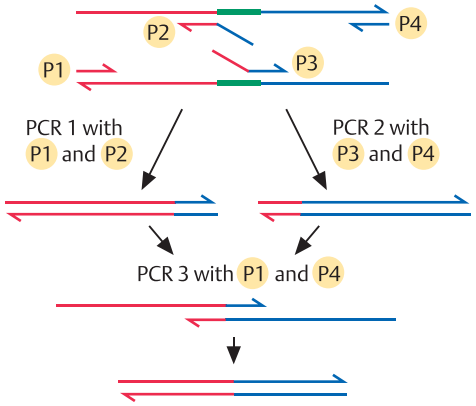
3 fusion of two DNA fragments



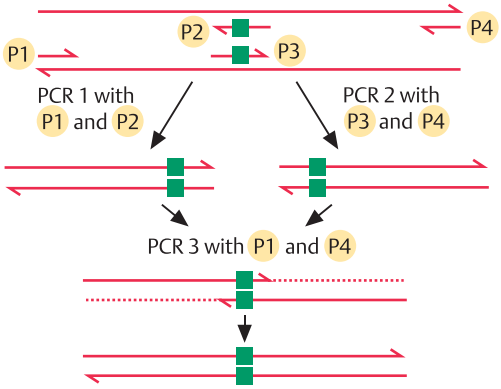
4 insertion of a DNA fragment



5 removal of a DNA segment



6 site-directed mutagenesis



DNA: synthesis and size determination

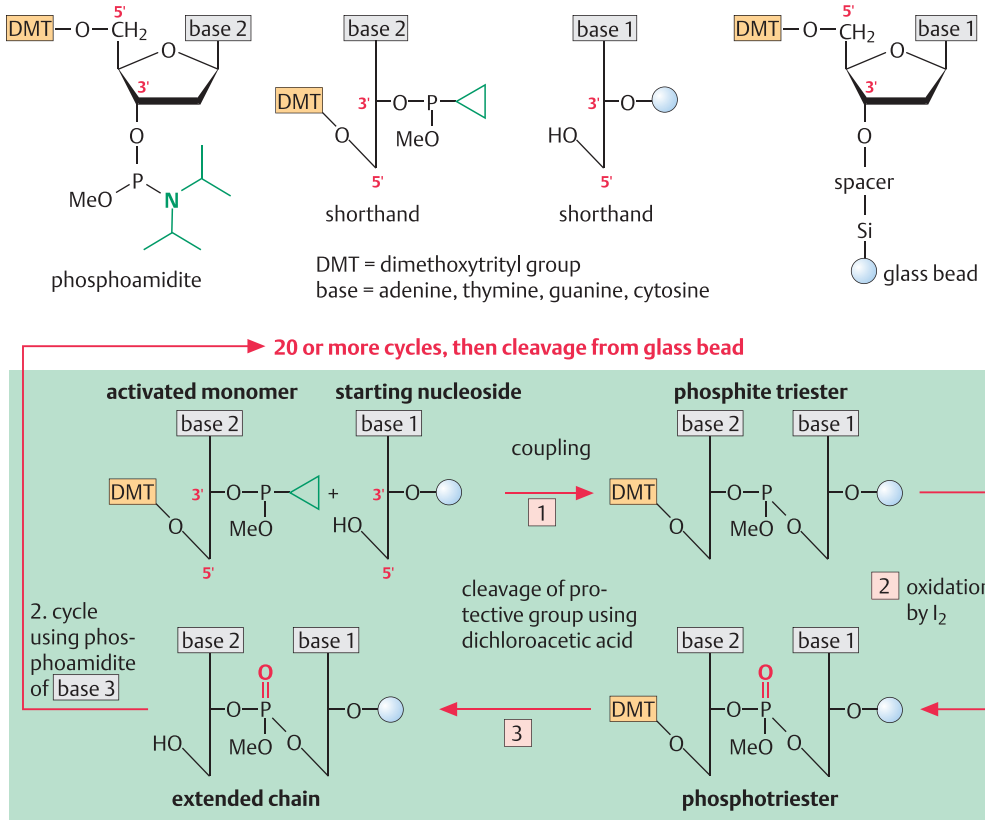
General. Short single-stranded DNA fragments up to ca. 100 bp (oligonucleotides) are chemicals that can be synthesized simply, quickly, and economically in the laboratory. They are useful for various steps in genetic engineering, e. g., as primers for PCR. For molar mass determination of DNA fragments up to ca. 30kbp, gel electrophoresis is used and standardized with DNA fragments of known M_R.

DNA synthesis. The method of choice is the phosphoamidite method, which is usually carried out in an automated synthesizer. All 4 nucleotide bases (A, C, G, and T) are present as phosphoamidites, in which the 3' phosphite group is protected by diisopropylamine and a methyl group. The 5' hydroxyl group of the deoxyribose and the amino groups of the purine and pyrimidine bases are also protected. First, nucleoside 1 is bound to an insoluble carrier material. Chemical unblocking of the 5' hydroxyl group leads to a nucleophilic attack on the tetrazol-activated phosphoamidite group of nucleotide 2. The resulting phosphotriester bond is now oxidized to a 5-valence phosphate ester, using iodine. This cycle, which in contrast to the biosynthesis of DNA proceeds from 3'→5', is repeated for each base. After the complete DNA fragment has been synthesized, all protecting groups are removed, and the single-stranded oligonucleotide is purified by gel electrophoresis or HPLC. Even if a 98 % yield is achieved in each reaction cycle, the total yield for a 20mer oligonucleotide is only 67 %, and only 45 % for a 40mer, resulting in DNA mixtures that are difficult to analyze and purify. For the synthesis of longer DNA segments or whole genes, complex strategies are necessary and are usually based on PCR. Oligonucleotides are mostly used 1) for synthesis of gene fragments or short genes, 2) as probes or primers for the identification or isolation of gene fragments from genomic or cDNA using hybridization or PCR, 3) for site-directed mutagenesis of a gene, and 4) for DNA sequencing. DNA synthesis is usually carried out by specialized laboratories that provide good quality and fast delivery at an acceptable price (2014: <0.30 US$/bp, delivery within days).

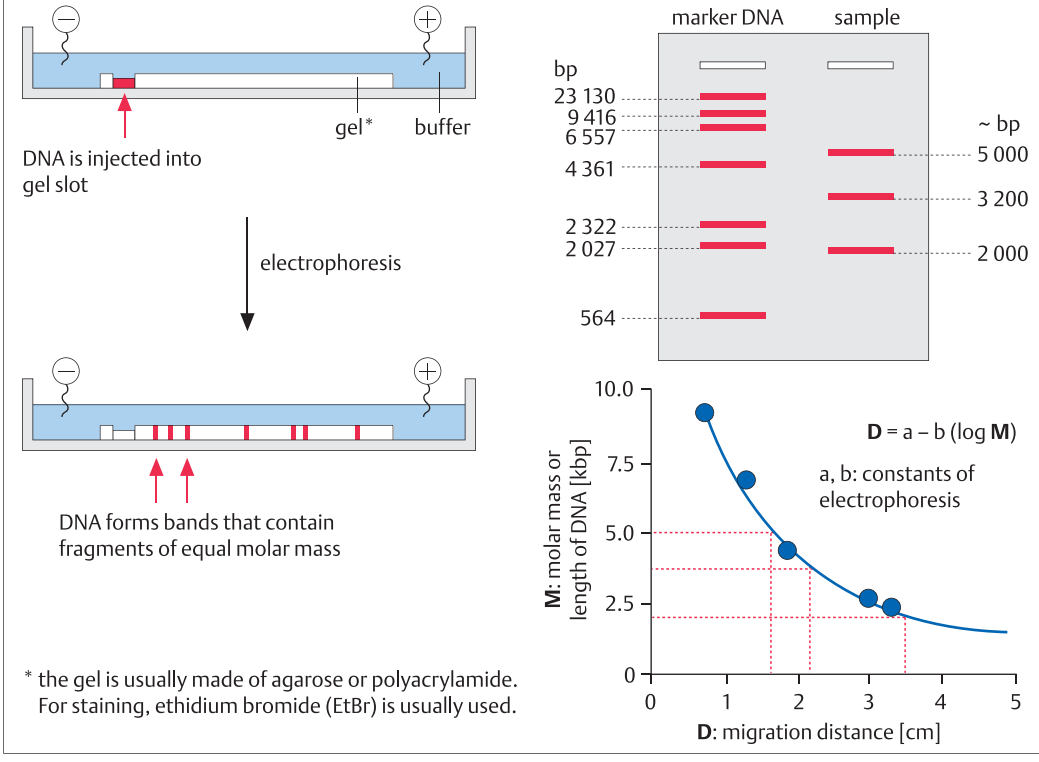
Size determination of DNA. Due to its net negative charge, DNA is easily separated by gel electrophoresis. Gels usually consist of agarose (large pore size), polyacrylamide (small pore size), or mixtures of both materials that allow one to define a mesh size that permits rapid analysis of the distribution of molar masses in a mixture of DNA fragments up to a size of ca. 30kbp with high precision. In most protocols, denaturing conditions are used (SDS-PAGE): if electrophoresis in a polyacrylamide gel (PAGE) is carried out in the presence of the surfactant sodium dodecyl sulfate (SDS), the mobility of single-stranded DNA depends only on its molecular mass, because the formation of secondary structures and intermolecular aggregates is prevented. Detection of DNA in a gel is done either by staining with ethidium bromide, by autoradiography, using radioactive labels, or by labeling with luminescence markers such as rhodamine:luminol (→84). Ethidium bromide is mostly used; however, it is genotoxic and must be used under appropriate safety conditions. The sensitivity of the ethidium bromide method is limited to >25ng DNA. DNA labeled with ^{32}P, using ^{32}P-labeled ATP and nick-translated with DNA polymerase I, can be detected at much lower concentrations, but requires radiation-safety equipment and routine monitoring for contamination. As a result, less demanding protocols based on fluorescent dyes are increasingly being used, e. g., SYBR Green, which is 25- to 100-fold more sensitive than ethidium bromide and allows a detection limit of >250pg DNA. Analysis is carried out in a phosphoimager after excitation with UV light. The M_R of a fragment can be calculated from its migration distance, but usually a set of DNA markers of various M_R is used for this purpose. Analysis of the molar mass of DNA may be important for restriction analysis of unknown DNA fragments, for constructing restriction maps, and for the identification of genes and gene fragments from chromosomal, plasmid, or viral DNA after PCR cloning.

Gene and genome synthesis. By combining oligonucleotide synthesis with PCR methods, complete genes, multi-gene metabolic pathways or genes for antibody libraries have been synthesized. The complete genome of *Mycobacterium capriolum* with 10^6 bp was synthesized in 2007 using such methods (→320).

Chemical solid-phase synthesis of DNA



Determination of DNA molar mass using gel electrophoresis



* the gel is usually made of agarose or polyacrylamide. For staining, ethidium bromide (EtBr) is usually used.

DNA sequencing

General. Two alternative methods can be used to sequence DNA: the Sanger-Coulson and the Maxam-Gilbert procedures. Both permit the sequencing of single-stranded DNA fragments up to a length of ca. 600bp. The sequence of longer DNA stretches must be derived from overlapping shorter fragments. When very long DNA fragments are sequenced, as in genome sequencing, highly automated methods are used. They rely on base-specific fluorescent dyes instead of the radioactive labels used in classical sequencing. Genome sequencing is highly demanding in terms of computer-based comparison of a very large number of sequences (an exercise in bioinformatics) (→324, 326).

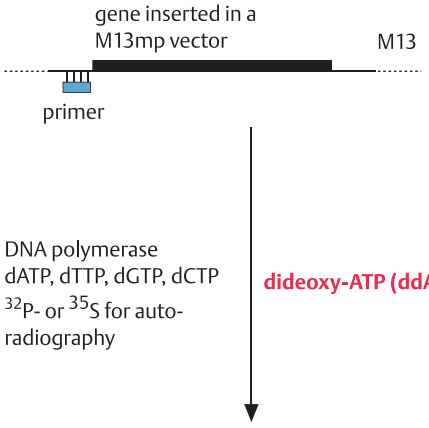
Sanger-Coulson method. DNA is cloned into an *E. coli* host infected with phage M13, resulting in phage progeny with single-stranded DNA. It serves as template for sequencing, using the Klenow fragment or, more frequently, T7-DNA polymerase, a short synthetic oligonucleotide as primer, and the 4 deoxynucleotides dATP, dTTP, dGTP, and dCTP as substrates. Double-stranded DNA is synthesized along the single-stranded DNA template. To four identical reaction vessels containing this reagent mixture, one of the four dideoxynucleotides (ddATP, ddTTP, ddGTP, or ddCTP) is added. Incorporation of these nucleotides at their complementary positions causes termination of DNA synthesis in a statistical manner, leading to a mixture of all possible DNA species terminating at each of these nucleotide analogs. Separation of this DNA mixture by gel electrophoresis allows the molar mass of the fragments to be identified, and, implicitly, the DNA sequence. Visualization on the gel is usually accomplished by autoradiography, after addition of a ^{32}P- or ^{35}S-labeled nucleotide to the reaction mixtures.

Maxam-Gilbert method. This procedure is used less today. It is based on the partial chemical hydrolysis of double-stranded DNA in 4 independent chemical reaction sequences, after the labeling of one terminus. Each base-specific reaction involves several steps (e. g., treatment with formic acid, dimethyl sulfate, hydrazine, etc.) and leads to (partially) selective cleavage at this base in the DNA strand, resulting in a family of terminally labeled DNA fragments which, as in the Sanger-Coulson method, are separated by gel electrophoresis and visualized by autoradiography. This method can also be run automatically, using a solid-phase procedure and labeling the terminal nucleotide with a fluorescent marker.

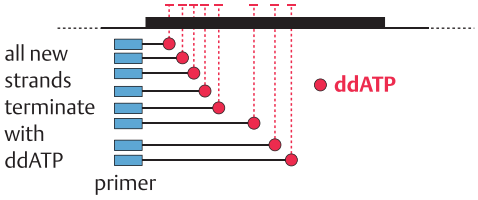
High-throughput sequencing. The long preferred method was based on the Sanger-Coulson procedure, with the following modifications: 1) double-stranded DNA can be sequenced by using specific primers in a PCR-type reaction (cycle sequencing), 2) the four dideoxynucleotides used for chain termination are labeled by coupling one of four different fluorescent markers to each base. This allowed all four nucleotides to be detected in a single reaction assay and, after time-resolved separation of the DNA fragments by flow-through gel electrophoresis, allowed determination of the molar mass of each fragment, leading directly to the DNA sequence. The read length was ca. 900bp, the duration of one cycle was 13h plus 2h of setup time. In commercially available instruments with 96 parallel electrophoresis lanes, the sequencing capacity was thus slightly less than 100,000 bases in 15h. If capillary electrophoresis was used instead of gel electrophoresis, the read length was reduced to ca. 650b, but the separation time was only 3h plus 1h of setup time. Thus, with a commercial capillary sequencer having 96 capillaries, 65,000 nucleotides were sequenced in 4h, or ca. 400,000 per day, in one instrument. Sequencers with 384 capillaries became also available. In the decade since this paragraph was written (2004), advances in high-throughput sequencing have been impressive and have made a strong impact on basic and applied molecular biology and on medicine (see chapter Megatrends, →312). For example, as a microbial genome sequence can now be solved within a day, a particular enzyme of a microbial organism may now be identified faster by using bioinformatic tools such as BLAST in combination with the sequenced genome as compared to standard cloning procedures. To correct for reading errors, genomes are sequenced several times, and data a corrected by alignment. To this end, "sequencing factories" have evolved that use high-performance robots for sample preparation, sequencers, and supercomputers for the final sequence alignment.

DNA sequencing after Sanger and Coulson

a annealing of the primer

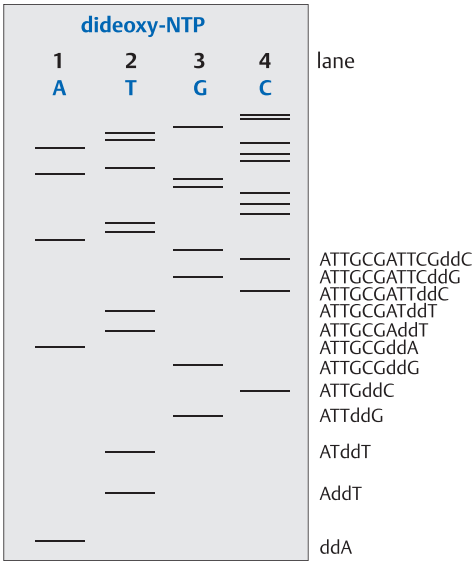


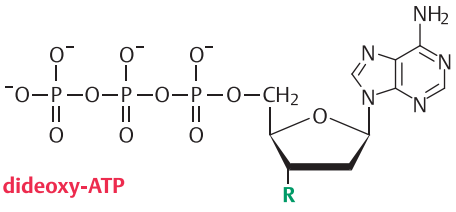
b example: synthesis of strands in lane 1
addition of dideoxy-ATP



ddTTP has been added in lane 2,
ddGTP in lane 3, and ddCTP in lane 4.

c gel electrophoresis and autoradiography





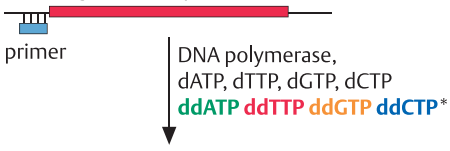
dideoxy-ATP

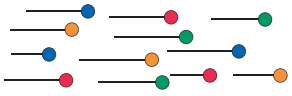
R = OH deoxy-ATP (dATP), normal nucleotide, strand synthesis

R = H dideoxy-ATP (ddATP), termination of strand elongation

High-throughput sequencing

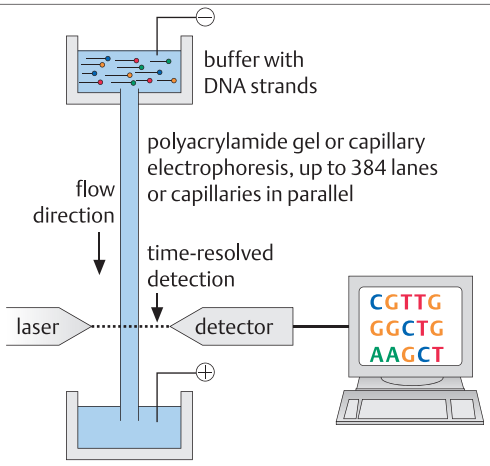
DNA single strand by PCR





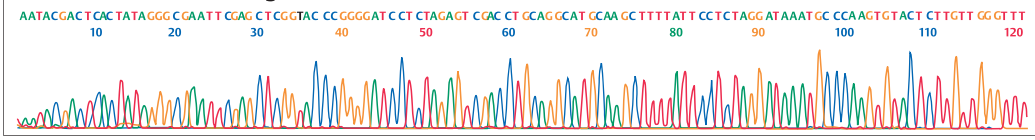
DNA strands of various chain lengths
with terminal fluorescence marker

separation by electrophoresis,
time-resolved detection



*using dideoxynucleotides with base-specific fluorescence marker for chain termination

time-resolved detector signal



Transfer of foreign DNA in living cells (transformation)

General. In nature, DNA is transferred into living cells in various ways: 1) transfer by plasmids, phages, or viruses (conjugation, transduction, transfection), or 2) direct uptake (transformation) (→8, 38). Cells that have incorporated foreign DNA are called transformed cells. In genetic engineering experiments, foreign (heterologous) DNA is usually transferred and expressed in a host cell. The transformation methods used are partially of biological and partially of technical origin.

Plasmids occur nearly exclusively in bacteria (→38). Most are circular double-stranded DNA molecules that replicate independently of the bacterial chromosome, but they can integrate into chromosomal DNA (episomes). Plasmids contain an origin of replication and, nearly always, one or several genes that are advantageous to the bacterium, e. g., a gene coding for antibiotic resistance. Plasmid DNA can be easily separated from chromosomal DNA and, like the latter, can be manipulated with enzymes in vitro. This has led to a wide range of cloning and expression vectors for use in genetic engineering. The most important functional properties of a plasmid vector are 1) an origin of replication (*ori*) for replication in the host organism; 2) optional: origins of replication for other host organisms (shuttle plasmids) – this allows construction of an appropriate vector in an easy-to-handle organism such as *E. coli* before transferring the genetic information to the desired, but more complex, host; 3) unique restriction sequences for inserting a gene only at the desired position (MCS, multiple cloning sites); 4) one or several resistance or auxotrophic markers for selection of positive recipient clones. Reporter genes facilitate the screening of transformed clones. For example, in the "blue-white" screening often used in combination with pUC plasmids, the plasmid-coded *lacZ'* gene complements the chromosomal *lacZ* gene of an *E. coli* host strain that lacks the *lacZ'* gene sequence (deletion *lacZΔM15*); only transformed *E. coli* clones can synthesize functional β-galactosidase, which in turn hydrolyzes the leuko dye 5-bromo-4-chloro-3-indolyl-β-D-galactopyrano-side (X-gal), forming dark blue 5,5'-dibromo-4,4'-dichloro indigo. Thus, transformed colonies are blue, whereas clones in which a foreign gene has been inserted into the multiple cloning site, the open reading frame coding for *lacZ'* is interrupted and the colonies remain white. Plasmid vectors are usually smaller than 10kbp, to facilitate manipulation and prevent their elimination from dividing cells due to negative selection pressure. Most plasmid vectors have been developed for *E. coli*, but plasmids useful in cloning experiments have also become available for Bacillus, Pseudomonas, Streptomyces, Lactobacillus, and some other bacteria. Plasmids are rare in eukaryotes. One of the few exceptions is the 2μ plasmid of *Saccharomyces cerevisiae* (→14). The Ti plasmid derived from the soil bacterium *Agrobacterium tumefaciens* has become an important vector for transforming dicotyledon plants (→280).

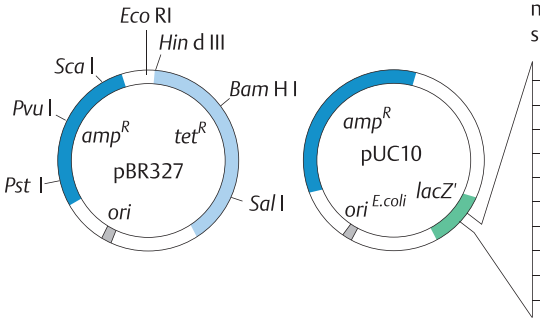
Bacteriophages and viruses permit the transfer of DNA into a host cell by transfection. Phage and viral vectors are attenuated by removing gene segments that are responsible for cell lysis or other mechanisms of pathogenicity. Specific phages are known for most bacterial species; many of them are used in genetic engineering, e. g., the λ and M13 phages (→8) for experimentation with *E. coli*. A small number of vectors based on attenuated viruses are also available for transforming plant, insect and animal cells.

Nonbiological methods comprise chemical and physical procedures. A method often used with *E. coli* is "heat shock transformation:" cells pretreated with $CaCl_2$ or $RbCl_2$ are made "competent" for the uptake of foreign DNA through a brief "heat shock" (42 °C for 60s). Electroporation uses a short electrical pulse that leads to the transient formation of pores in the cell membrane, resulting in uptake of the DNA. Biolistics (→280) are mentioned in the section on transformation of plant cells. For transformation of animal cells or plant protoplasts (both contain no cell wall), DNA can be precipitated as the Ca salt on the surface of the cells, initiating endocytosis. Other procedures include fusion of cells with liposomes containing DNA (lipofection) and microinjection of DNA into the nucleus of eukaryotic cells (→266). With these methods, the number of transformed cells remains small, and the procedure must be optimized for each experimental setup.

Plasmids

	example	size [kbp]	occurrence
f(ertility) plasmids	F-plasmid	95	*Escherichia coli*
r(esistance) plasmids	RP4	54	*Pseudomonas* sp.
toxin (colicin) plasmids	ColE1	6.4	*Escherichia coli*
degradative plasmids	TOL	117	*Pseudomonas putida*
virulence plasmids	Ti plasmid	213	*Agrobacterium tumefaciens*

Cloning vectors



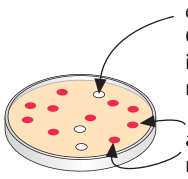
multiple cloning site (MCS)

- *Hind* III
- *Sph* I
- *Pst* I
- *Sal* I, *Acc* I, *Hinc* II
- *Xba* I
- *Bam* H I
- *Sma* I, *Xma* I
- *Kpn* I
- *Sac* I
- *Eco* R I

two cloning vectors for *E. coli*

$ori^{E.coli}$	origin of replication for *E. coli*
amp^R	selection marker 1 (ampicillin resistance)
tet^R	selection marker 2 (tetracycline resistance)
lacZ'	gene fragment of β-galactosidase

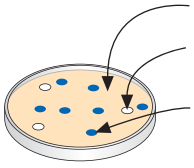
screening for antibiotic resistance



$amp^R tet^S$-recombinant clone. Gene fragment was inserted into tet^R using BamH1 or SalI restriction sites

$amp^R tet^R$: no recombinant clone

blue-white screening



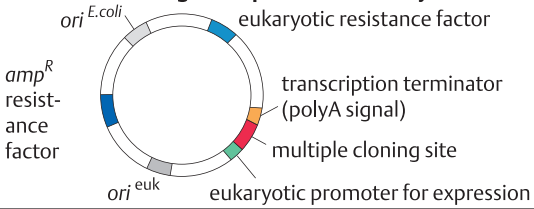
agar + X-gal + IPTG + ampicillin

white colony = recombinant clone (*lacZ'* destroyed)

blue colony = no recombinant clone (*lacZ'* intact)

X-gal: chromogenic substrate for β-galactosidase
IPTG : inducer for β-galactosidase promoter

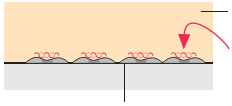
shuttle vector for gene expression in eukaryotes



$ori^{E.coli}$	origin of replication for *E. coli*
amp^R	selection marker for *E. coli*
ori^{euk}	origin of replication for eukaryotes (e.g., of 2µ plasmid for *S. cerevisiae*, of SV40 virus for animal cells)

Transformations based on nonbiological methods

endocytosis



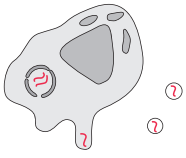
calcium phosphate solution

DNA is deposited on the cell surface

monolayer of animal cells

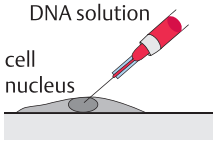
liposome fusion (lipofection)

transfer of DNA into the cell nucleus

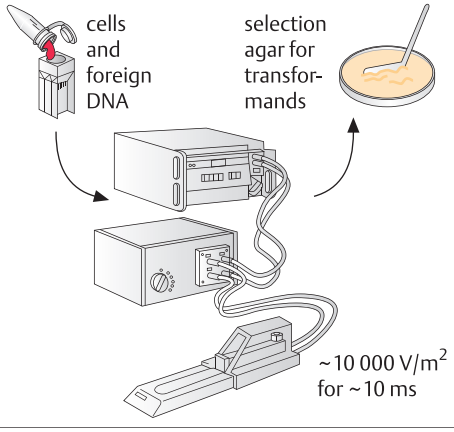


fused liposome (DNA in red)

microinjection



electroporation



Gene cloning and identification

General. PCR methods are widely used for cloning genes whose sequences are known or can be derived from a protein's partial amino acid sequence (→52). If a gene of unknown protein or gene sequence must be cloned, a genomic library (prokaryotes) or cDNA library (eukaryotes) is prepared and deposited in a set of transformed host organisms (usually *E. coli*). The desired gene must be present in this library if its product is functionally expressed and can be identified by its activity (shotgun cloning), through its transcribed mRNA (but only if it is a major cell product), or by immunological identification of the gene product (Western blot).

Cloning with PCR methods. Once enough sequence information about the desired gene or gene product is available, synthetic primers can be constructed, which allows the gene to be cloned from DNA or cDNA using PCR. Degenerate primers must be used if the precise sequence of the gene is unknown, e. g., if it is derived from an amino acid sequence. Often, PCR cloning is combined with the insertion of a restriction site for later ligation into an expression vector.

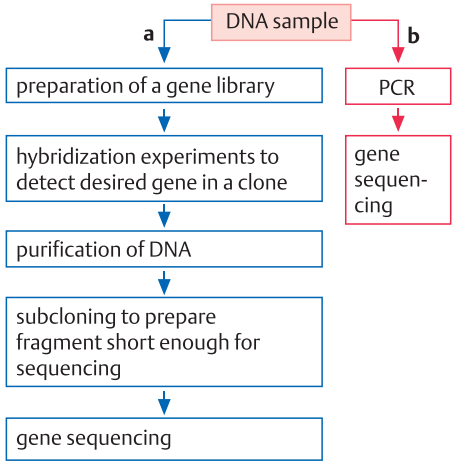
Preparation of gene libraries. DNA or cDNA from the donor cells (bacteria, plant, mammalian, or insect cells) is digested by restriction nucleases, and competent host cells (usually *E. coli*) are transformed with this library (→68). Because very many different clones must be analyzed for heterologous gene inserts, efficient selection methods are crucial. Thus, the vectors used for transformation may contain marker genes that confer resistance to antibiotics, e. g., amp^R or tet^R, so that only transformed cells will survive on the selection agar. In the marker-rescue approach, the auxotrophic mutant of a wild-type strain is transformed with DNA fragments from a gene library containing the gene responsible for the auxotrophic properties: the transformants that are complemented by this gene can grow on minimal medium without additives. Usually two marker genes are used: one for the selection of transformants, the other as a component of the multiple cloning site (MCS) into which the foreign DNA is inserted. The successful integration of a foreign DNA into the MCS can then be recognized by a loss of the marker phenotype, e. g., antibiotic resistance. After these preparatory steps, a gene library can now be analyzed for transformants containing the desired gene.

Detection of genes and gene products. The most important procedures are based on 1) gene-specific hybridization of a DNA or RNA probe, and 2) expression of the gene product. In the first procedure, a gene-specific probe complementary to a sequence in the desired DNA is synthesized and radioactively or otherwise labeled; it is then used for hybridization experiments with single-stranded DNA obtained from the transformants (Southern blot), or directly in colony or phage hybridization. mRNA transcribed by a transformant can be analyzed in an analogous way, using a labeled DNA or RNA probe (Northern blot). If the DNA sequence of the gene product is not known, a Western blot can be used. Here, the gene library is prepared by using an expression vector. The desired protein may then be found with an immune reaction using labeled antibodies. Even protein fragments may be discovered by this technique – a useful property, since the encoding gene may have been cut during preparation of the gene library, being distributed among two or more transformants. If a gene library is to be searched for regulatory elements, e. g., for promoters, vectors that contain a reporter gene (e. g., for luciferase or green fluorescent protein) behind the MCS are used. A promoter isolated from the gene library is then detected by the expression of the reporter gene.

Other detection methods. PCR methods based on specific primers are usually used to monitor the insertion of a gene into chromosomal DNA. Restriction patterns may provide the first indication of the successful cloning of a new gene, if novel restriction patterns appear during gel electrophoresis upon comparison of wild-type and transformant genes.

Identification of gene functions. The first step is usually based on bioinformatics: by comparing a number of partial sequences of the isolated gene with a sequence database through a BLAST analysis (*Basic Local Alignment Search Tool*), the most probable functions of an isolated gene are ranked. BLAST takes into account mutation probabilities between amino acids and thus is also useful to establish phylogenetic relationships (→326).

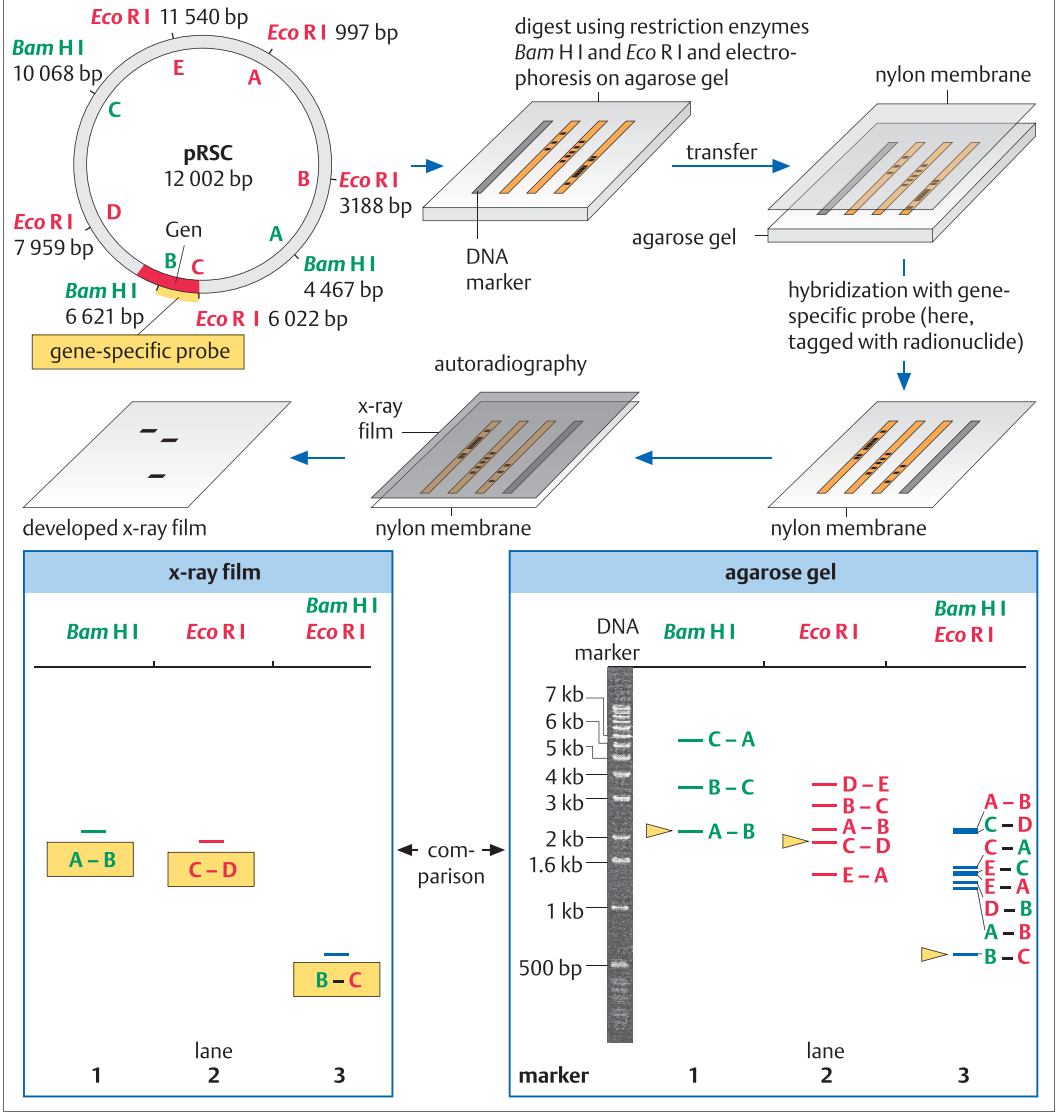
Cloning of genes



Detection methods

Southern blot	detection of DNA	hybridization with labeled DNA or RNA probe
Northern blot	detection of mRNA	hybridization with labeled DNA or RNA probe
Western blot	detection of proteins	immunological detection by labeled antibodies
reporter groups	detection of regulatory elements	expression of genes coding for reporter proteins

Southern blot



Gene expression

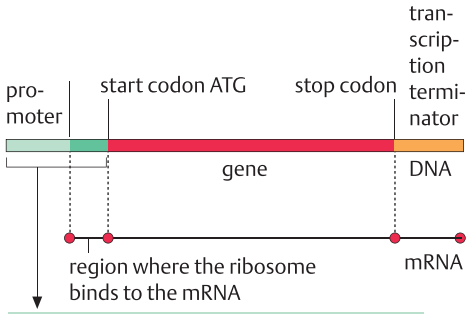
General. A main objective of genetic engineering is to express a foreign gene or operon (several coordinated genes) in a host organism. Foreign linear DNA can be integrated into a host organism *via* homologous recombination through homologous, double strand DNA sequences (the standard in all higher eukaryotic cells). For gene expression in prokaryotic organisms such as *E. coli*, homologous recombination is also done, but the use of expression vectors, usually a plasmid that replicates outside of the chromosome, is preferred. As it can be lost during cell divisions, it usually bears an additional antibiotic resistance gene, and fermentation must be done in the presence of the respective antibiotic. In expression vectors, the foreign gene is often preceded by an inducible promoter, allowing the gene to be switched on or off by appropriate external conditions. In higher organisms, it has become possible to also direct and express foreign DNA in a desired compartment, e. g., the chloroplast, by using appropriate leader sequences.

Expression vectors for prokaryotes. A typical expression vector for bacteria contains an origin of replication (*ori*), a marker gene to enable selection of transformed clones, and the foreign structural gene or operon (ORF = open reading frame), with its start codon ATG and its terminal stop codon. Several recognition sequences provide the appropriate commands to the transcription and translation machinery of the cell to form the gene product. In *E. coli*, RNA polymerase binds to sequences upstream of the ORF (the so-called –35 and –10 boxes) and transcribes DNA into mRNA. Transcription ends at a transcription terminator region that is located downstream of the ORF, sometimes by forming a stem-loop region in the mRNA. For constructing expression vectors, inducible promoters are usually preferred. For example, the promoter of the lactose operon (*lac*) of *E. coli* can be switched on by adding the inducer isopropyl-β-D-thiogalactoside (IPTG) to the medium; in the presence of this inducer, a repressor protein is removed from an operator sequence, allowing for RNA polymerase to bind to the promoter and to transcribe the gene into mRNA. Many expression vectors are commercially available. They usually contain multiple cloning sites (MCS).

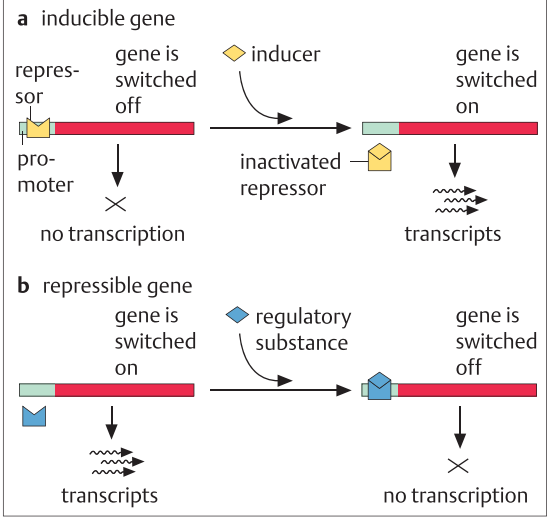
Expression vectors for eukaryotes have a similar structure. They contain a selection marker, often an inducible promoter with consensus sequences (TATA, CCAAT, and GC boxes), a start codon (ATG), followed by a multiple cloning site, and a terminator sequence. The mRNA obtained during transcription in an eukaryotic cell is polyadenylated at its 3'-position (polyA tail) and carries a 7-methylguanosine triphosphate residue at its 5' position ("cap"). Specific signal sequences may lead to expression of the gene product in a desired cellular compartment. Eukaryotic expression vectors for higher organisms rarely replicate autonomously: they are usually inserted into a chromosome of the host organism by recombination. To select clones of transformed animal cells having many copies of the heterologous gene, auxiliary genes are cloned into the vector that provide the cells with resistance to toxic culture-medium components. Thus, coexpression of high copy numbers of dihydrofolate reductase (DHFR) or neomycin phosphotransferase (→98) with the desired gene product ensures that only cells transformed with the desired gene can survive in a medium that contains high concentrations of methotrexate or neomycin.

Promoters. Promoters are characterized as strong or weak, and tight or loose. Promoters used in technical processes should be strong and tight, i. e., they should remain switched off in the absence of inducer. Typical promoters for *E. coli*-based processes are the *lac*, *trp*, *tac* or *rha*-promoters, which are induced by adding a reagent to the medium (e. g., L-rhamnose in the case of the *rhaPBAD*-promoter). In contrast, the λP_L-promoter is induced by raising the temperature of the medium from 30 to 42 °C. For expressing foreign genes in fungi and yeasts, the GAL10 galactose promoter is often used in *Saccharomyces cerevisiae*, the alcohol oxidase promoter AOX in *Pichia pastoris*, and the glucoamylase promoter in *Aspergillus* strains. The metallothionein promoter (→98) is popular for expression in animal cells. For transgenic plants and animals, promoters that are regulated by the host organism are often used. Thus, in transgenic animals the target gene is often cloned behind the strong lactalbumen promoter of the mammary gland (→272); the recombinant protein is then formed in large quantities after induced lactation.

Expression vectors



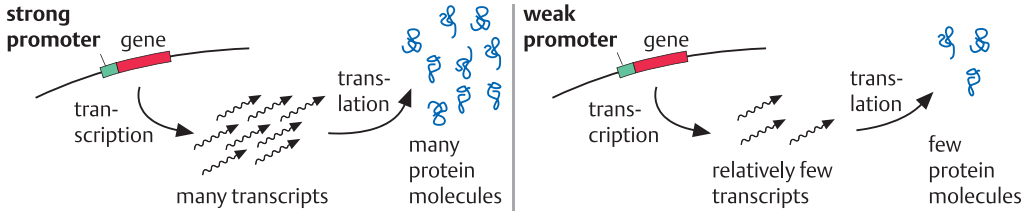
Induction and repression



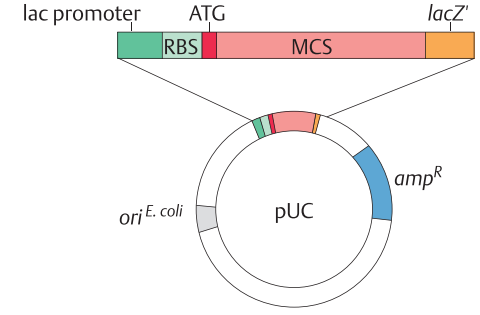
Inducible promoters (examples)

promoter	protein	inducer	host organism
lacZ	β-galactosidase	IPTG	*E. coli*
λP_L	–	temperature rise 30 – 42 °C	*E. coli*
GAL10	β-galactosidase	galactose	*S. cerevisiae*
AOX	alcohol oxidase	methanol	*Pichia pastoris*
metallothionein	metallothionein	Zn^{2+}	animal cells

Promoters

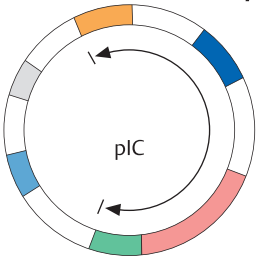


Expression cassette for *E. coli*



RBS ribosome binding site
ATG start codon
lacZ' sequence for blue-white screening
MCS multiple cloning site with recognition sequences for restriction enzymes
ori E. coli selection marker
origin of replication for *E. coli*

Expression cassette for *Pichia pastoris*



- 3' end of the alcohol oxidase gene
- origin of replication for *E. coli*
- selection marker amp^R for *E. coli*
- alcohol oxidase promoter
- multiple cloning site with recognition sites for restriction enzymes
- His4 auxotrophic marker for *Pichia pastoris*
- vector integrates into *Pichia pastoris* chromosome

Gene silencing

General. The targeted silencing of genes is an important technique in basic molecular biology studies and in biotechnology, for example, to eliminate undesired properties in the breeding of domestic animals and plants, in the development of microbial strains, and in medicine, e. g., for tumor therapy. Unlike random mutagenesis based on chemical mutagens or radiation, genetic techniques have the potential to knock out specific genes. Experimentally, this is done by recombination or RNA-based techniques. While gene silencing has undoubtedly been successful in some cases, most phenotypes are multigenic, and it is usually very difficult, often impossible, to assign a desired phenotype to the function of a single gene. Based on endonucleases, a powerful new method was developed for the multipleed editing of whole genomes, repairing or removing genes or whole operons.

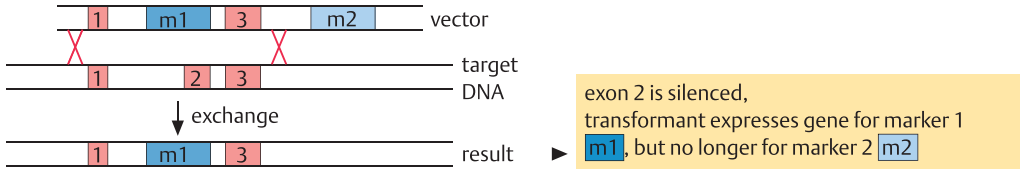
Knockout by recombination. DNA replacement vectors used for creating knockout mutants are nearly homologous to the gene or exon to be silenced, but contain a mutation or deletion that results in a nonfunctional gene product after the vector has recombined with the chromosomal DNA. Reliable recombination is ensured only if the length of the inserted fragment exceeds ca. 150bp. Because recombination events are rare ($< 10^{-3}$), markers are required to select transformed cells. Growth inhibitors are often used for this purpose, e. g., methotrexate, whose inhibitory activity can be overcome by cells that express enough dihydrofolate reductase (DHFR) (→98).

RNA-based techniques. In RNA-based techniques, the gene to be inhibited is ligated in the reverse direction into an expression vector that is used for transformation of the host cell. The mRNA that is formed during transcription of this gene (antisense RNA, asRNA) is complementary to the mRNA of the normal gene and prevents synthesis of the gene product. Probably both species form an RNA double strand that either is not bound to the ribosomes or is rapidly degraded by RNases (→42). The antisense technique offers an interesting new concept in medical therapy, which complements gene therapy: if a disease is not due to an erroneous gene product, but to its excessive formation, replacement of the gene is less promising than interference with its translation by antisense techniques. In fact, the cancer-inducing properties of glioblastoma cells in brain tumors, which are due to errors in the formation of insulin-like growth factor, were decreased by expression of asRNA. The as RNA was expressed by means of an episomal vector containing a metallothionein promoter. Whereas RNA interference is based on the formation of double-stranded small interfering RNA (siRNA), which forms a RISC complex with the target mRNA that leads to the degradation of the latter, a different concept for using asRNA in medicine is based on its direct injection; to this end, RNA analogues such as phosphothionates must be used because RNA is quickly degraded *in vivo* by RNases. In 2012, over 40 as-oligonucleotides and siRNAs were being clinically evaluated. As of 2014, two of them were registered in the USA: Fomivirsen, as a treatment for cytomegalovirus infection in immune-suppressed patients, and Mipomirsen, a drug that inhibits formation of apolipoprotein B (treatment of homozygous or heterozygous familial hypercholesterolemia). An early successful example of a fruit produced by antisense techniques was the FlavrSavr™ tomato, whose fruits can be left to ripen and form aroma on the vine and have a long shelf life (→282). Plants transformed by such vectors contain marker genes, e. g., coding for antibiotic resistance. Critics have pointed out that horizontal gene transfer of these resistance markers may lead to the spread of antibiotic resistance throughout the ecological system.

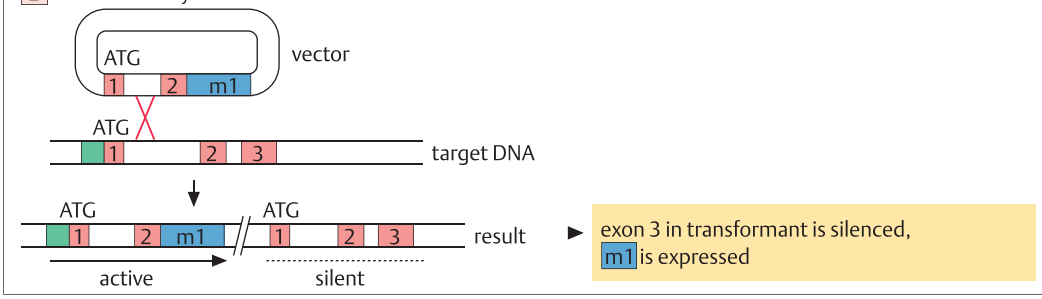
Genome editing. Various methods allow for the targeted editing of genomes. Among them, homologous recombination using double- or single-stranded DNA is quite efficient. Most of these methods use protein engineered DNA-binding nucleases, zinc finger nucleases (ZFNs) or transcription activator-like effector nucleases (TALEN) The RNA-dependent CRISP/Cas9 nuclease system also allows for the multiplexed engineering of large genomes. These methods have a great potential for gene therapy as well as in plant in animal breeding (→264, 274). Thus, it was applied to the repair of an impaired dystrophin gene in iPS cells of a patient suffering from Duchenne muscular dystrophy. Also, the biosynthesis of the toxic lectin ricin could be disrupted in seeds of the oil plant *Ricinus communis*.

Knock-out of genes by gene targeting

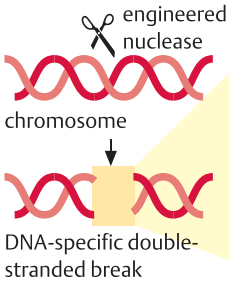
1 positive–negative selection



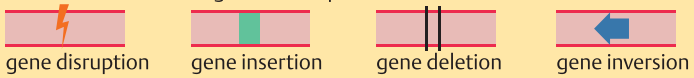
2 inactivation by insertion



Nuclease-mediated genome editing



non-homologous end joining
without added homologous DNA repairs with indets



homology-directed repair
with added homologous DNA, insert/replace DNA

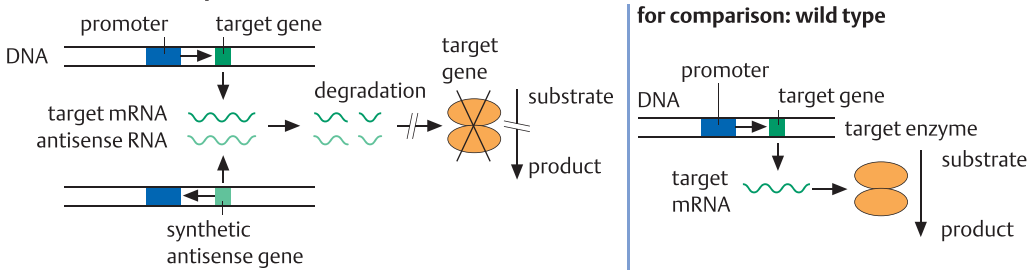
four families of engineered nucleases

1 engineered mega-nucleases and re-engineered homing endonucleases
2 transcription activator-like effector nucleases
3 zinc-finger nucleases
4 CRISPR-Cas system (**c**lustered **r**egularly **i**nterspaced **s**hort **p**alindromic **r**epeats)

Eukaryotic selection markers

marker	cell type	indicator
recessive adenosine deaminase	CHO mutant	9-β-xylofuranoxyl
dominant dihydrofolate reductase	all	methotrexate
neomycin phosphotransferase thymidine kinase fusion protein	all	neomycin sulfate
metallothionein I	all	Cd^{2+}, Zn^{2+}

Antisense technique



Epigenetics

General. The term epigenetics is used to describe heritable and non-heritable changes in gene function that cannot be explained by changes in DNA sequence. Epigenomics describe such changes through the entire genome. In higher organisms, cell division is accompanied by differentiation into various types of cells. In this epigenetic process, covalent changes at the DNA or the histone level occur, leading to a silencing of DNA regions that henceforward defines the cell's functional identity throughout its life cycle. DNA damage (man: 10,000 events per cell per day), though largely repaired, may also lead to an accumulation of epigenetic changes over time.

Mechanisms. Epigenetic changes are caused by a) post-translational modifications of histone amino acids, and b) methylation of DNA, mostly at GpC sites. Histone modifications may include acetylation, methylation, ubiquitinylation, phosphorylation or sumoylation (modification by a SUMO1 protein), changing the shape and/or the charge of a histone protein domain which wraps a particular gene sequence, leading to changes in gene expression. A wide range of enzymes have been described that carry out such modifications, such as histone lysine methyltransferase (KMT) which acts on histones H3 and H4. There are also enzymes that demethylate histones such as histone lysine demethylase (KDM), which removes up to three methyl groups from histones H3 or H4. Several histone modifications may occur simultaneously and work together to regulate gene transcription ("histone code"). DNA methylation (→48) occurs frequently in repeated sequences of GpC sites. The product, 5-methylcytosine, can spontaneously deaminate to thymidine, increasing the probability of permanent mutations. As a consequence, CpG sites are frequently mutated and rare in the genome, The enzymes involved in these modifications, DNA methyltransferases, occur in replication sites, and their elimination is lethal in experiments with transgenic mice. Hemimethylation of only one strand of DNA is used to silence one X DNA strand during meiosis (epigenetic templating). In mitosis, the DNA methyltransferase, DNMT1, transfers methylation patterns to each newly synthesized strand after DNA replication ("maintenance transferase"). There is evidence that small non-coding RNAs participate in epigenetic regulation as well, possibly for modulation of promoters during gene expression.

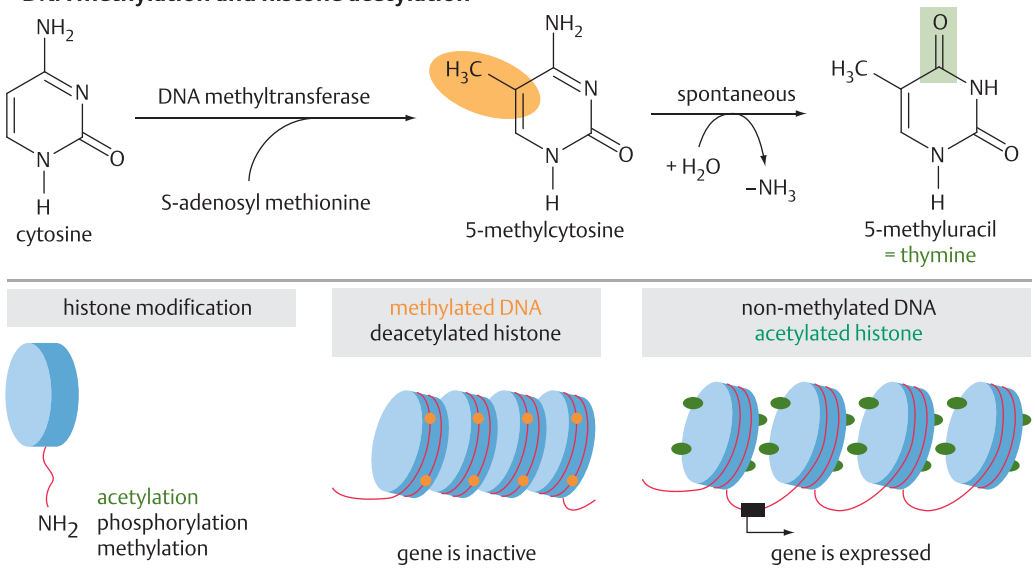
Functions. Epigenetics are often divided into predetermined and probabilistic epigenesis. Somatic epigenetic inheritance is predetermined as a key to cellular differentiation of multicellular eukaryotic organisms. Examples for probabilistic genetics is the inheritance of traits by genomic imprinting (father and mother carry different epigenetic patterns in their germ cells). In mice, it was shown that traumatic experiences were passed on via epigenetics through two generations (progeny of mice which had received foot shocks during exposure to a cherry blossom odor rejected this odor much more strongly than untrained controls). Transgenerational inheritance in men was shown in a study on three generations with > 300 probands in the Swedish village Överkalix where paternal grandsons of adolescent men exposed to famine in the 19th century died less of cardiovascular diseases but more from diabetes in a surplus food environment.

Relevance to biotechnology. Compared to healthy cells, tumor cells often show different methylation patterns. Thus, the diagnosis of specific mutations in the tumor repressor gene *BRCA2* (for breast cancer 2, early onset) provides a risk assessment for breast cancer or a hereditary breast-ovarian cancer syndrome. The BRCA2 protein is part of a DNA repair complex, and epigenetic changes in the *BRCA2* gene may lead to reduced functionality.

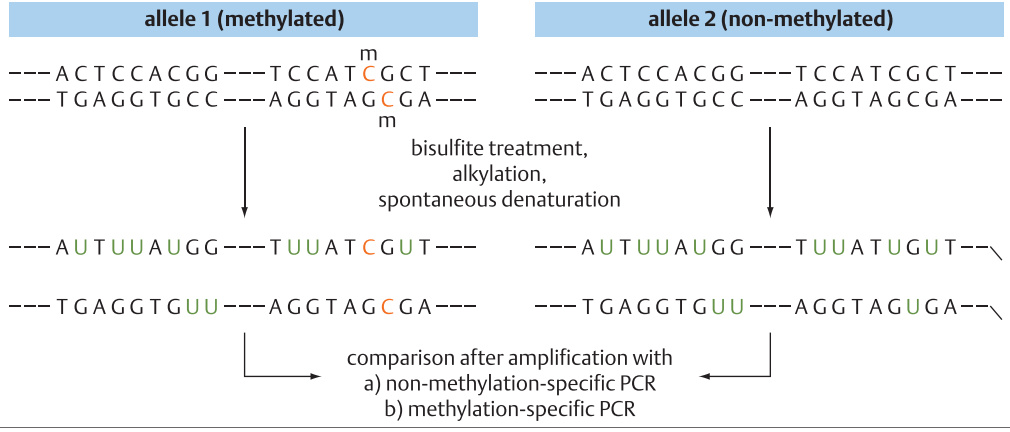
Microorganisms. Bacteria methylate DNA adenine instead of cytosine. This process is important for proof-reading of freshly replicated DNA, but also in the control of virulence in pathogenic microorganisms such as Salmonella, Vibrio, Yersinia and Brucella. Thus, by genome sequencing of *E. coli* O104:H4, which had caused some 50 deaths in Germany in 2011, it was discovered that this strain had not only acquired a shiga toxin, but also a DNA-methyltransferase from a phage that had completely modified the *E. coli* "methylome" and led to high virulence.

Analytical tools. Bacterial genome methylation can be assayed by special PCR protocols (→52) or by single-molecule real-time sequencing (SMRT). A classical procedure is the pretreatment of DNA with sodium hydrogensulfite. This transforms non-methylated cytidines to uracils. By comparing the sequences before and after treatment, methylation sites can easily be detected.

DNA methylation and histone acetylation

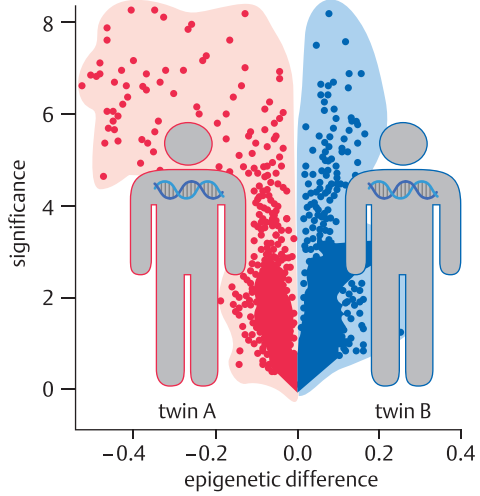


Analysis of DNA methylation: distinguishing bisulfite-generated polymorphisms



Monozygous twins

are epigenetically different



Food behavior

may lead to epigenetic differentiation

Isogenic Avy mice of same age and sex

Coat color and obesity depend upon the epigenetic status of the Avy allele

Gene libraries and gene mapping

General. Even the small genomes of phages and viruses are much too large for direct sequencing of their DNA or RNA. Thus, genomic DNA is cut into fragments and cloned into vectors, stepwise approaching the size of fragments that can be sequenced. After fragments have been sequenced, the sequence information is analyzed by computer and made into a complete DNA- or RNA-sequence map of the genome (physical map). Due to the redundancy of identical base sequences, larger genome sequences can only be considered correct if enough markers have been identified (gene mapping). Of particular practical importance as markers are sequence-tagged sites (STS).

Gene libraries comprise a collection of DNA fragments that, together, constitute the complete genome. They are prepared by cutting genomic DNA into smaller DNA pieces, using ultrasound or restriction enzymes (→46), and inserting the fragments into vectors. To prepare large fragments, restriction enzymes that cut at rare sequences are preferred, e. g., *Not*I, which recognizes 5'-GCGGCCGC-3', occurring statistically only every $4^8 = 65{,}536$bp. The occurrence of a sequence in any genome depends strongly, however, upon the GC content of the DNA and on the presence of repeating DNA sequences.

Vectors. For preparing gene maps and for sequencing, genomic DNA fragments are cloned into vectors, by which they can be easily amplified through the transformation of host cells and from which they can be re-isolated. The size of the genomic DNA determines the number of clones that are needed to create a complete gene library. For the larger eukaryotic genomes, the vectors of choice are "artificial chromosomes" such as yeast artificial chromosomes, YACs (→14), for the construction of gene libraries in *Saccharomyces cerevisiae*, or bacterial artificial chromosomes, BACs (→20), for gene libraries in *E. coli*; they allow DNA inserts up to 300kbp to be packaged. For sequencing, these fragments are still much too large. Subcloning is usually carried out with λ-derived vectors, e. g., cosmids (→8), which allow 30–45kbp of foreign DNA to be packaged.

Gene mapping. The classical method of gene mapping relies on observing coupled phenotypes. For example, gene maps of Ascomycetes such as *Neurospora crassa* and *Saccharomyces cerevisiae* were obtained long ago by tetrad analysis. Because all the daughter cells are found in one ascus in the same sequence in which they were formed during meiosis, the spores can be easily isolated and analyzed for their phenotypic properties. Molecular genetics has strongly expanded these methods. For smaller genomes, restriction fragments can be generated and sequenced, often leading to direct physical mapping of a gene's position. By using well-chosen probes for PCR or DNA hybridization, one can also obtain genetic markers for larger genomes. An important procedure in this context is fluorescence in-situ hybridization (FISH) (→84) of a large DNA sequence. High resolution (within 10 kbp) can be obtained if this method is combined with "DNA combing". For this, a polylysine-covered slide is dipped into a solution containing large DNA fragments. If the slide is slowly ($0.3\ \text{mm}\ \text{s}^{-1}$) pulled out of the solution, the DNA fragments align in parallel, offering perfect conditions for hybridization with a marker.

STS (sequence tagged sites) are DNA sequences 100–500 bp long, which occur just once in the whole genome (→296). This implies that STSs do not include sequences found in repetitive DNA. STSs are often obtained from clone libraries containing large genome fragments, e. g. from a YAC or BAC library. If large eukaryotic genomes must be analyzed, chromosome-specific libraries are often used. Individual chromosomes can be isolated after staining with a fluorescent dye and sorting by flow cytometry (fluorescence-activated cell sorting, FACS), since the amount of dye that a chromosome binds depends on its size. After a collection of STS has been found for a given genome, it is simple to find out, with appropriate PCR primers, if they have a neighboring or far-distant position in the genome: if they are close together, a collection of overlapping gene fragments from a gene library should yield additional hybridizing gene fragments carrying the same STS. STS markers thus are excellent "mapping reagents" for molecular-coupling analysis of gene segments.

Gene libraries

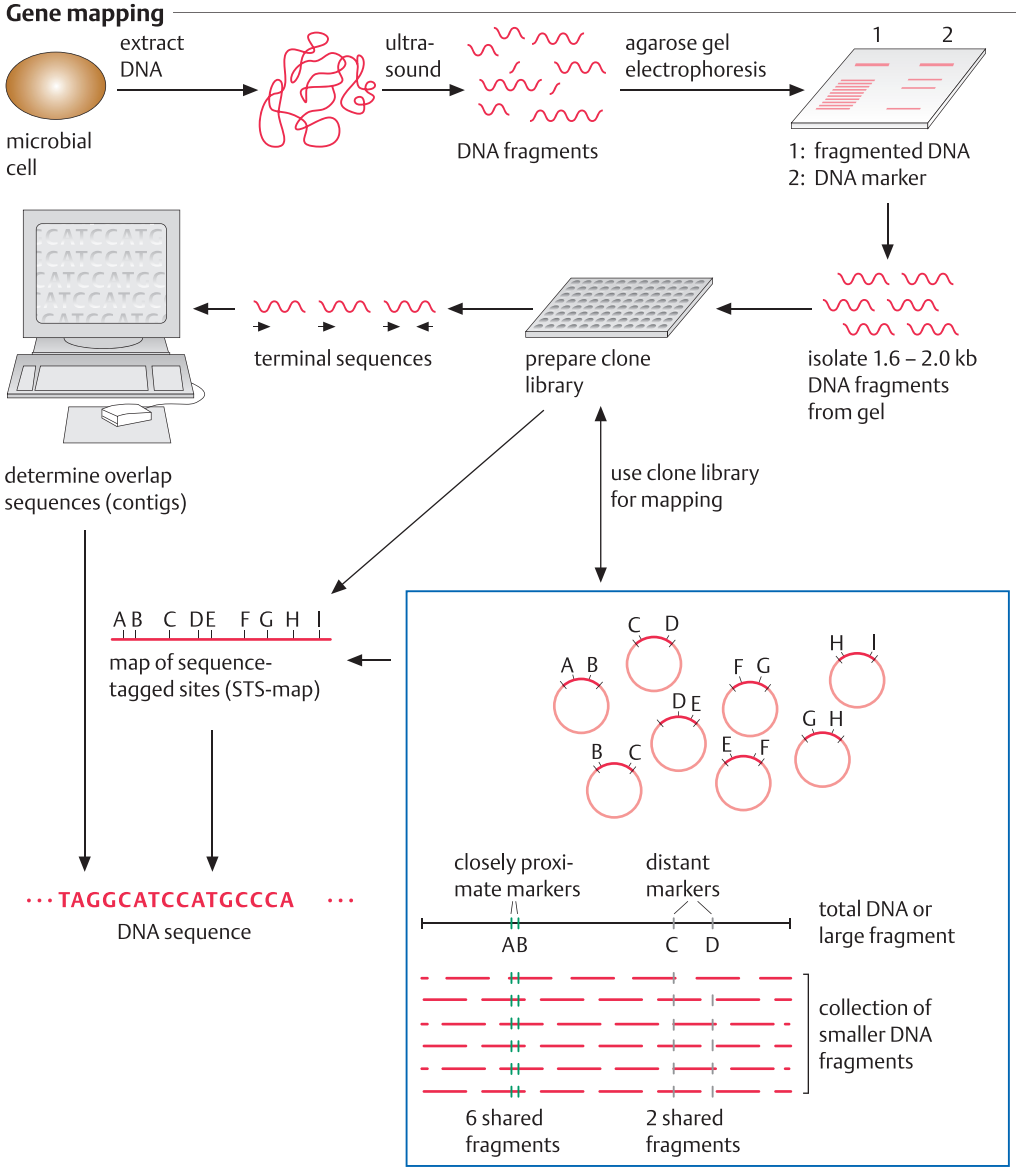
organism	number of required clones (P = 95 %) (haploid genome) b	λ-vector (EMBL4) a ≈ 17 kbp	cosmid a ≈ 35 kbp	BAC a ≈ 250 kbp	Y AC a ≈ 1000 kbp
E. coli	4 800 000	850	410	56	13
S. cerevisiae	14 000 000	2 500	1 200	167	41
Drosophila melanogaster	170 000 000	30 000	14 500	2 036	508
tomato	700 000 000	123 500	59 000	8 387	2 096
man	3 000 000 000	529 000	257 000	35 948	8 986
frog	23 000 000 000	4 053 000	1 969 000	275 602	68 901

$$N = \frac{\ln(1 - P)}{\ln(1 - a/b)}$$

P = probability

a = average length of DNA inserts in the vector (bp)
b = total size of the genome (bp)

Gene mapping



Genetic maps of prokaryotes

General. Genetic maps of microorganisms have been prepared by observing changes in the phenotype after conjugation (transfer of DNA from a donor to a recipient cell), after transduction (transfer of DNA pieces among bacteria by a phage), and after transformation (uptake of naked or plasmid DNA). Physical genome maps (the complete DNA sequence of a genome) have existed since 1995. They have been prepared by clone contig mapping or the shotgun procedure.

Genetic maps. Many changes in phenotype can be observed easily and rapidly in bacteria. Thus, loss of the capacity to form spores or flagella, or introduction of antibiotic resistance, can be used as a phenotypic marker. For the elucidation of metabolic pathways, blocked mutants, which have lost the ability to carry out one or more steps in a pathway have been studied for a long time (→24); by adding the precursor molecule, this loss can be overcome. The short generation time of many bacteria (< 1h) is a bonus for the microbial geneticist. When observation of two phenotype modifications in a recipient cell is combined with measurement of the time period required for conjugation or transduction, the distance of the genes coding for the two phenotypes can be estimated (linkage analysis). Prokaryotic gene maps are therefore measured in minutes or centisomes (n/100 of chromosome length 100). The time required for complete transfer of the *E. coli* genome into a recipient cell is 100 min at 37 °C.

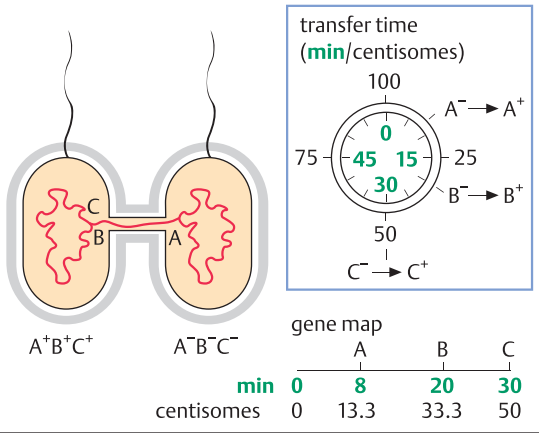
Physical genome maps: clone contig maps. An important task in genome sequencing is to identify those clones in a genome library that contain neighboring DNA sequences (contiguous = neighboring). Clone "contigs" for a whole genome can be combined by clone fingerprinting techniques. Thus, the occurrence of overlapping restriction maps or sequence-tagged sites (STS) in two clones indicates that they contain overlapping parts of the genome sequence and can be used to deduce its overall sequence with the aid of computer programs. Wherever possible, phenotype-related markers from genetic maps are also used. For positional cloning of a gene from a nearby marker, chromosome walking is often used. In this approach, labeled RNA from the starting clone can be prepared for hybridization experiments or primers are constructed from the terminal sequence of a clone and used to prepare a PCR product from the contiguous clone, which is suspected to contain the desired gene. The DNA sequence of the complete genome is identified step by step using the incomplete genetic map as a template.

Physical genome mapping: shotgun method. This time-saving method is based on the concept that any DNA sequence of 600bp can be directly and rapidly determined. Originally, genomic DNA was cut by restriction enzymes. Today, ultrasound or hydrodynamic procedures ("nebulizer") (→312) are preferred for DNA fragmentation, providing small overlapping fragments, whose end-sequences are determined. A supercomputer is then used to compile the genome sequence from the sequenced overlaps (assembly of sequenced fragments to longer contigs). Due to the power of high-throughput sequencing (HTS), the efficiency of modern teraflop computers, and the huge backup of thousands of prokaryotic genomes which have been solved, this procedure allows a typical microbial genome (1–5Mbp) to be mapped within days, depending on the type of sequencer available. Because genetic maps containing marker sequences are often available (when sequencing of the *E. coli* genome began in 1990, there were only 1400 markers, corresponding to a mean distance between markers of 3300bp on the 4.64Mbp genome), the computed results obtained by shotgun sequencing can continuously be validated. By comparing genomes and their functional units, a unified view about the metabolic modules involved in life is emerging.

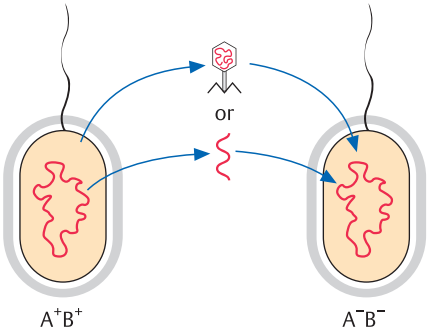
Bioinformatics. The computer programs used in genome projects are aimed mainly at reliable determination of sequence homologies. Even rare sequencing errors (99 % precision) lead to errors in practically all recorded gene sequences and render it necessary to compare results from multiple sequencing experiments. Annotation is done using standardized procedures, e. g., using a "Prokaryotic Genome Automatic Annotation Pipeline." orthology analysis helps to identify genes found in different species that originate from a common ancestor, as xenology is useful to identify genes, which have been horizontally transferred between two organisms.

Genetic maps from conjugation experiments

sequential transfer of marker genes A, B, C during conjugation

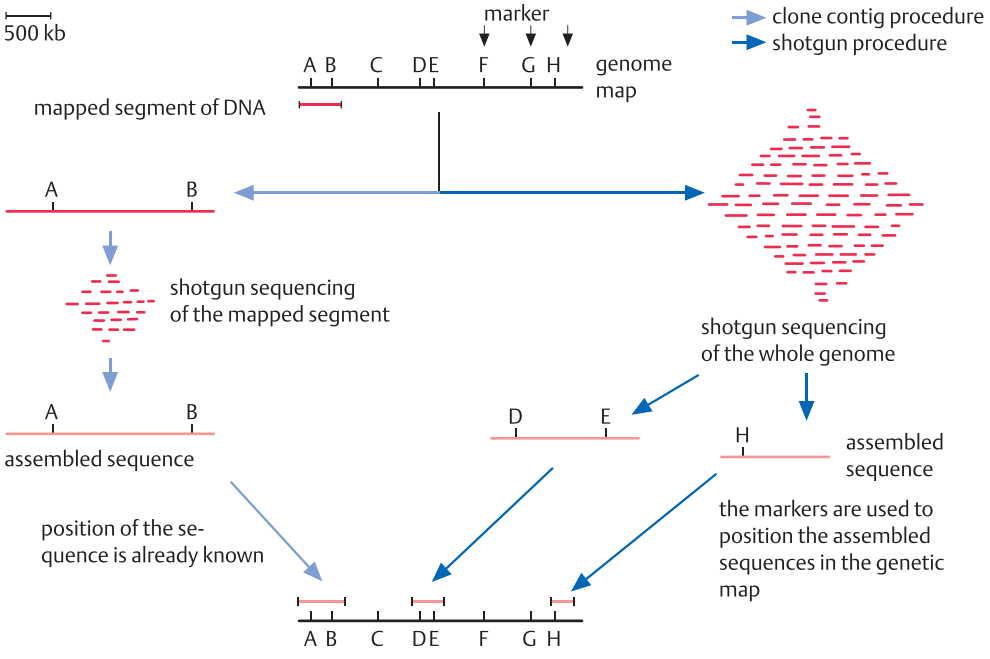


cotransfer of closely proximate marker genes A, B during transduction or transformation

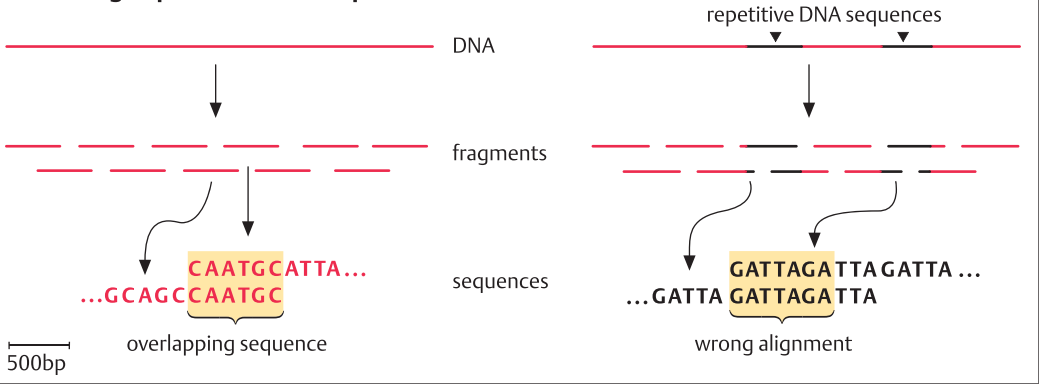


the frequency of $A^-B^- \rightarrow A^+B^+$ depends on the proximity of A and B on the chromosome

Sequencing of prokaryotic genomes



The shotgun procedure and its problems



Genetic maps of eukaryotes

General. Genetic maps of eukaryotes are produced, similar to prokaryotic maps, by coupling analysis of genetic phenotypes. Due to the diploid or polyploid set of chromosomes in eukaryotes, however, the gene sequence responsible for the phenotype may originate from different genotypes (heterozygous), which segregate during meiosis. Another feature of eukaryotic genomes is their much larger size, usually several billion base pairs. In addition, they contain introns and a significant number of repetitive DNA sequences, which hinder the search for unique sequences. In spite of these difficulties, the genomes of hundred of eukaryotic organisms have already been completed and the trend is now towards projects in "GigaScience," e. g., the exploration of the genetic basis of human diversity (human diversity project, 1,000 human genome project), complex medical inquiries (human cancer genome project) (→298), animal and plant breeding programs based on genomic information, or the million micro-ecosystems genome project.

Genetic maps. In experimentally accessible animals, genetic maps are based on pedigrees combined with linkage analysis: on observations of how phenotypic properties are linked during sexual reproduction, i. e., by meiotic crossing over (see genetics textbooks). Two phenotypes corresponding to genes that are close together on a chromosome are co-inherited more frequently than two phenotypes due to genes that are farther apart. In consequence, the recombination frequency of coupled phenotypes leads to a virtual genetic map, whose dimension is the percentage frequency of recombination. This classical method is complemented today by many molecular genetics methods. The fingerprints of different relevant DNA fragments can be determined and compared by restriction mapping, and primers tagged with fluorescent markers can be used to locate genes within large DNA fragments or chromosomes by fluorescence in-situ hybridization (FISH) (→84).

Genome sequencing. The genomes of higher organisms contain repetitive DNA sequences (satellite DNA, Alu sequences, retrotransposons, etc.) (→296, 298), which make unambiguous localization of a sequence within the overall genome difficult. Thus, SINE (short interspersed nuclear elements) ranging from 100 to 500 bp each, e. g., the Alu sequences, which constitute up to 20 % of a mammalian genome, and LINE (long interspersed nuclear elements), 6000–7000 bp long, account for up to 10 % of a mammalian genome. In addition, mini- and micro-satellite DNAs contribute another 5 %. In humans, microsatellite DNA consists of 10–50 copies of a very short-sequence repetition such as AC or ACCC, which occurs > 10,000 times and is distributed over the whole genome. Since each individual has a unique distribution of microsatellites, they are excellent genetic markers, e. g., in forensic investigations (→302), and also in breeding domestic animals (→268). Due to the high redundancy of repetitive sequences, genome sequencing by the shotgun approach, which is so useful for prokaryotic genomes, meets with considerable difficulty. As a consequence, contig sequencing of overlapping clones is widely used, combined with genome walking, the use of sequence-tagged sites (STSs) (→70) and of expressed sequences tags (ESTs) (→296) as markers. STSs and ESTs provide complementary information. STSs span the whole genome, but do not discriminate between coding and noncoding regions and may include repetitive sequences. ESTs are short sequences derived from cDNA clones. Because cDNA is prepared by converting spliced mRNA into double-stranded DNA with reverse transcriptase, the sequences of ESTs contain no repetitive sequences and each EST has a unique sequence. If primers derived from EST sequences are hybridized with a genomic library, all introns and repetitive sequences are undetected, but clones that contain complete or fragments of expressed genes can be identified. As a consequence, STSs and ESTs complement each other very well as mapping reagents. Once the clones of a genome library have been correlated with a physical map of the genome by one or more of the above methods, subcloning of the cosmid (→8), YAC (→14), or BAC (→20) clones in λ phage libraries ensues, followed by DNA sequencing (→68). Overlapping sequences are then analyzed with sequence-contig software to generate a complete sequence of the DNA of single chromosomes and, finally, of the whole genome. This computed map is validated with information from genetic maps.

Genetic maps

genes

m small wings

v shimmering eyes

w wide eyes

y yellow body

resulting position on a genetic map:

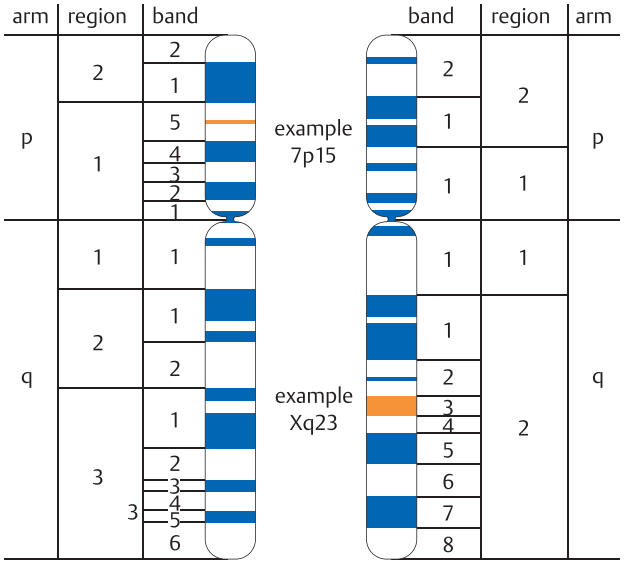
y	w	v	m
0	1.3	30.7	33.7

recombination frequency:

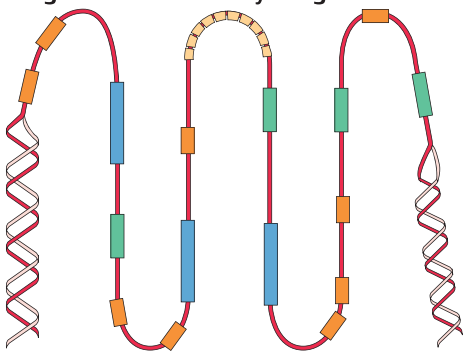
between m and v	3.0 %
between m and y	33.7 %
between v and w	29.4 %
between w and y	1.3 %

Chromosome nomenclature

chromosomes in the metaphase after Giemsa staining



Organization of eukaryotic genes



satellite DNA:
telomere and centromere region, minisatellites (16 – 64 bp), microsatellites (2 – 4 bp, up to 50 copies)

SINE (short interspersed repetitive elements), 100 – 500 bp

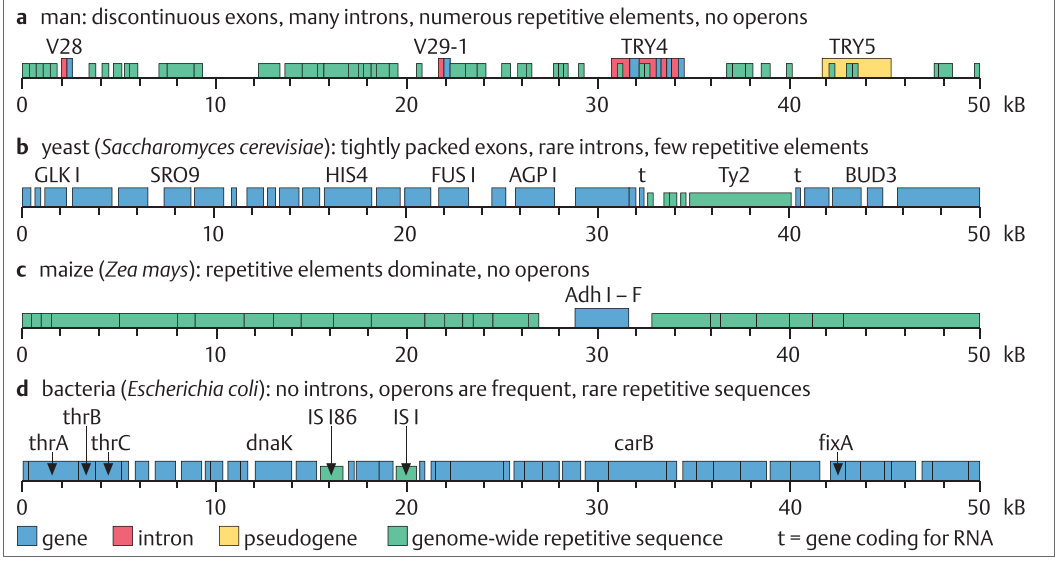
LINE (long interspersed repetitive elements), up to 7 000 bp

gene

centromere

telomere (5'-TTAGGG-3')

Gene structures of different organisms



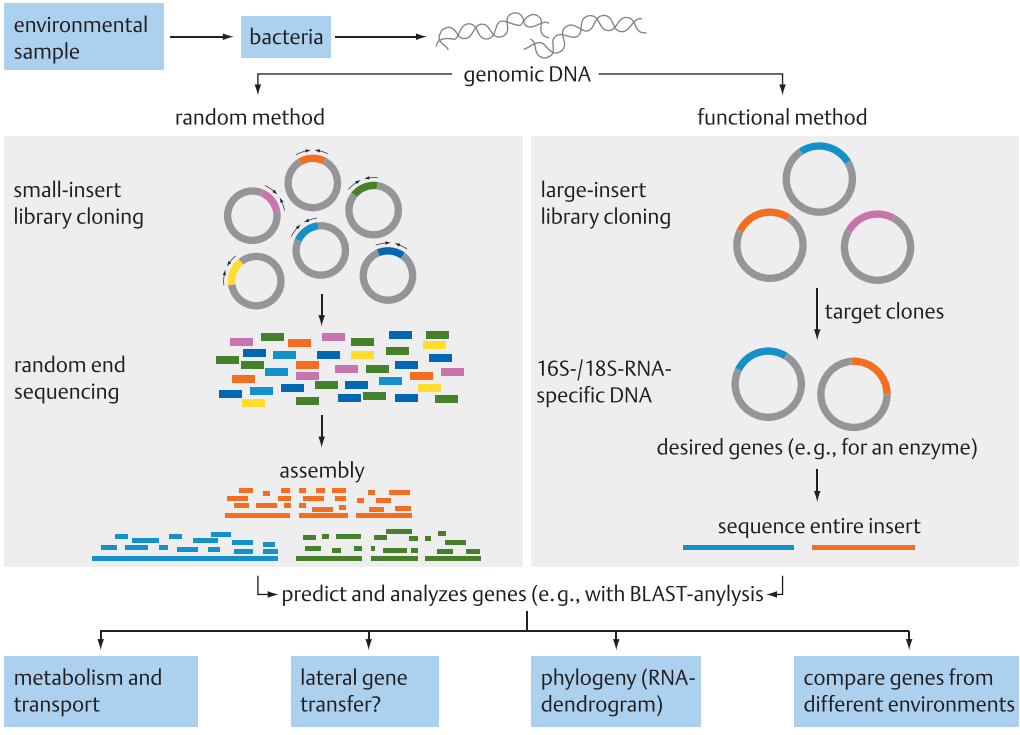
Metagenomics

General. The term metagenome has been coined to describe the sum of all genomic information contained in a biotope, a symbiotic community ("biocenosis"). Even larger contiguous communities with similar climatic conditions are termed a "biome." By genetic mapping of marker genes of defined species in a biome, e. g., of their mitochondrial DNA ("DNA barcoding") a taxonomic reference database is attempted to be generated. Microbial communities occur in sewage sludge (→286, 288), in the intestines of man and animals ("microbiome") (→118), but also in nutrient-rich soil and water samples. Most of the microorganisms living in such communities can not yet be cultivated. Due to the advances in high-throughput DNA sequencing (HTS) and bioinformatics, their gene pool has now become available and can be used for the cloning and expression of new enzymes or gene cassettes.

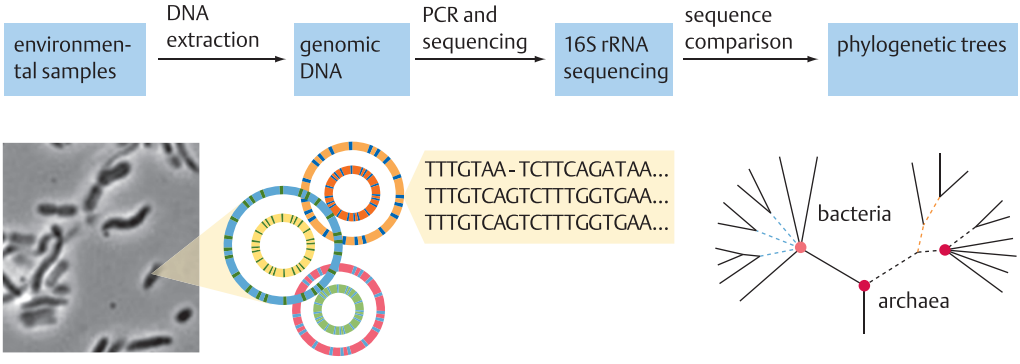
Methods. For the construction of a metagenomic library, e. g., from a sludge-, soil- or stool sample, DNA is extracted from the sample and degraded into DNA-fragments using restriction endonucleases (→46). The fragments are either directly sequenced or expressed in a host organism, usually *E. coli*. For a *functional approach,* gene products are then expressed and subjected to simple tests. Thus, lipases and esterases can be detected by the formation of clear halos when the host organism is cultivated on a turbid tributyrin-containing agar plate, or amylases by lucid halos on a brown agar that contains an iodine-starch complex. Active clones are then propagated, the foreign DNA, usually inserted in plasmids, is isolated and sequenced and the sequence is analyzed by bioinformatic procedures such as BLAST. Usually, many of the DNA fragments discovered in this way code for novel enzymes. A second, even more versatile procedure is in the *sequencing approach*. In the first step, one attempts to build a comprehensive gene bank containing all of the metagenomic DNA. Shotgun sequencing then leads to a large sequence space, which contains information about the diversity of the organisms present in the probe through sequence information for ribosomal 16S-RNA (prokaryotes) and 18S-RNA (eukaryotes). This information is mostly used to create an RNA-dendrogram. By comparison of DNA sequences to a DNA database, usually GenBank, a comparison of homologies to known DNA sequences allows for the targeted cloning of desired proteins. Gene cassettes that code, e. g., for the synthesis of antibiotics, can also be identified using a similar approach. In a few cases, this procedure allowed to elucidate the complete genome sequence of a non-cultivatable microorganism. Thus, the genome of *Buchnera aphidicola*, an endosymbiont of aphids, was reconstructed from bacterial DNA collected from this insect.

Applications. The knowledge about symbiotic communities and their interactions has advanced greatly in recent years and improved our understanding of processes occurring, e. g., in the human digestive tract. Through international consortia, the "microbiomes" (→118) of the oral cavity, the human intestines, the vagina etc. were sequenced and analyzed depending on age, diseases, food preference, etc. Sludge microbiomes have been another interesting target; not only could changes in populations over time and performance be studied, but plasmid metagenomes from sludges have revealed a high level of antibiotic resistance genes and mobile genetic elements, indicating a hot-spot for the formation of microorganisms which are cross-resistant to several antibiotics (→204). Another milestone of metagenome research is the *Global Ocean Sampling Expedition* of the Craig Venter Institute, which started in 2003. In this program, 100L samples of seawater was taken every 200 nautical miles, and metagenomes where analyzed. As early as 2004, this approach provided 1.6Gb of mostly new genetic sequences (submitted to GenBank), and pointed to the presence of some 2,000 species and >1 million potentially translated proteins, many of which were unknown. In the meantime, the research vessel, *Sorcerer II*, used for this expedition, has taken samples in all seven seas. At present (2014), it is on mission in the Amazonas region. Metagenomics is also a very useful approach for industrial applications. An early application (2000) was the isolation and propagation of a microbial strain (Thauera) that produced exopolysaccharide that had excellent sequestration properties for heavy metals. The German company B.R.A.I.N. keeps in its BioArchive large collections of culture collections in addition to DNA libraries, comprising some 150 million genes that have been prospected from various microbial habitats and whose application in screening have led already to several novel, patentable technical enzymes or "synthetic" microbial pathways, e. g., for the industrial production of antibiotics or vitamins.

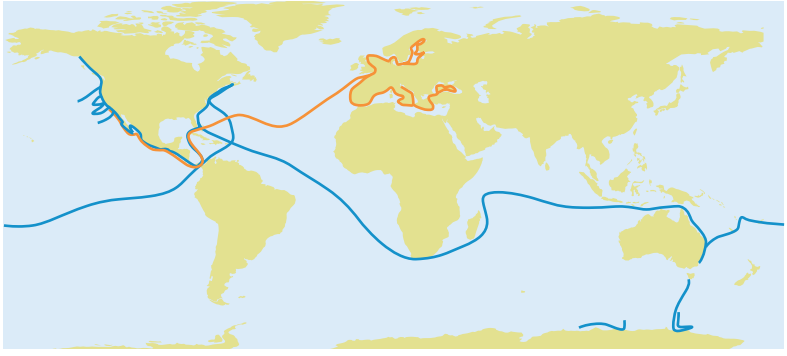
Metagenome analysis



Diversity of a metagenome: RNA dendrogram



Global Ocean Sampling Expeditions, J. Craig Venter Institute, La Jolla, USA



During just one expedition from Halifax through the tropical East Pacific (2007/2008), analysis of 6.3 billion bp led to the discovery of 6.12 million new genes and 1700 new protein families. Genetic diversity within one and among two sampling sites (distance: 200 nautical miles) was very high.

Cell biology

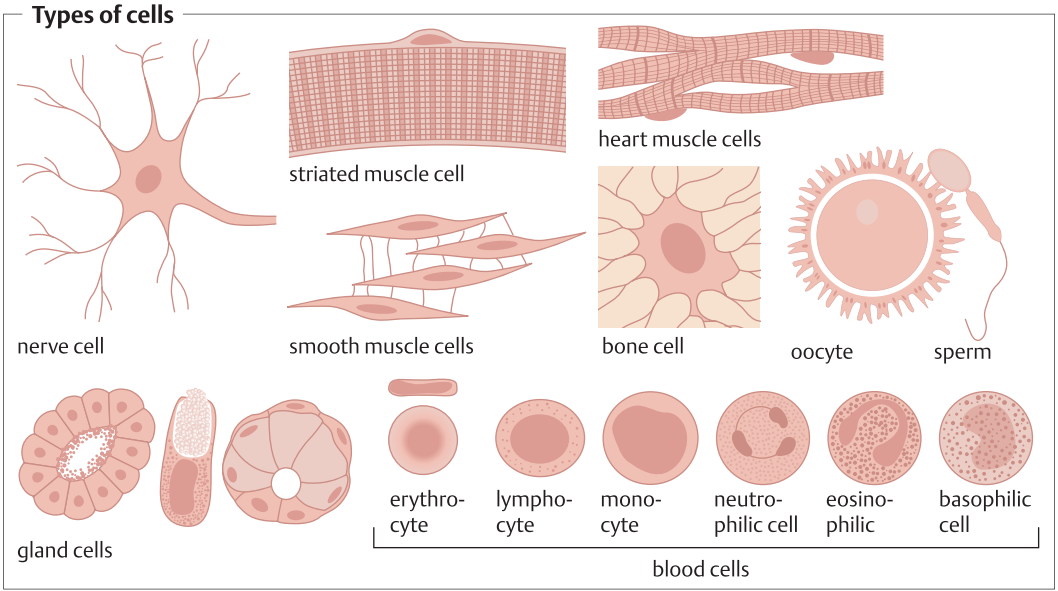
General. Cell biology has become a quite comprehensive discipline, and only a very brief introduction is possible in this pocket guide. In view of biotechnology, the focus is on the eukaryotic cell of multicellular organisms and man, with some emphasis on the human immune system (→80) and stem cells (→78).

Cells, tissue, organs. The principle components of a single eukaryotic cell, and elements of its biochemistry, are described in other parts of this book, though important aspects internal cellular organization such as the cytoskeleton, the many aspects of energy supply and cellular vesicle trafficking could not be outlined. Additional aspects found in multicellular organisms such as the cell cycle during mitosis, or the mechanisms for targeted destruction of the body's own cells by apoptosis, cannot be discussed in the framework of an introduction to biotechnology. Against the background of stem cell research and the success of biopharmaceuticals which relate to blood cells and the immune system, a short introduction to these latter topics will be provided. The organization and the interactions of the some 10^{14} cells in a human body is obviously quite complex, e. g., in view of cell differentiation or concerning the transport of substances and signals.

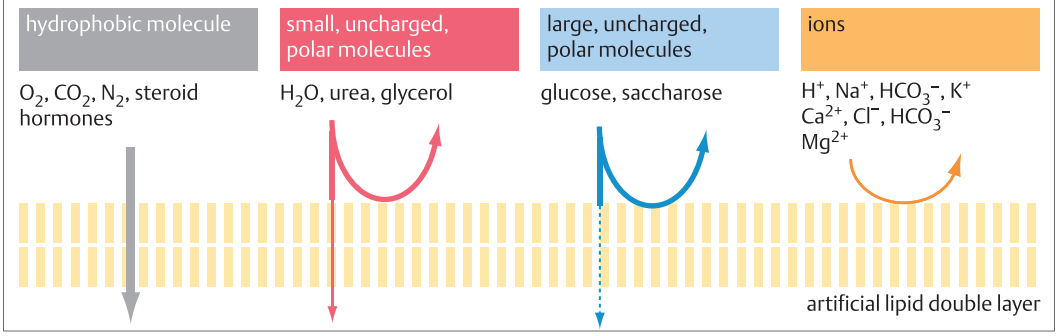
Cell differentiation. During the development of a fertilized oocyte to an embryo and ultimately to a complete organism (ontogenesis), specialized cells are formed through a process termed maturation. In the case of humans, about 200 different cell types are formed (fibroblast, myoblast, osteoblast, erythrocyte, neuron, etc.). Cells of the same type form tissues. To this end, they either generate an extracellular matrix from proteins and polysaccharides (for example, a connective tissue), or, through their cytoskeleton, a cell matrix (for example, an epithelium). The correct choice of cells required for a tissue is made by the cadherins, a family of membrane proteins. Tissues of various types combine to form organs such as, e. g., the epidermis, the sensory epithelia, the alveoli of the lung or the intestinal mucosa. In each tissue, stem cells are stored. They are not completely differentiated, can divide without limits, and differentiate on specific signals. As a consequence, they regenerate their particular tissue by continuously supplying fresh, differentiated cells.

Transport of substances and signals. Most substances have poor permeability across the lipid double membrane (→34). Only when membranes are endowed with channels and transporters, formed from membrane-associated or membrane-spanning proteins, are they able to import and export ions and molecules. Most transport processes require energy, e. g., via hydrolysis of ATP (ATP-driven pumps) (→26). For sending signals to cells, higher organisms utilize several hundred types of signal molecules. Examples for signal molecules are specific proteins, peptides, amino acids, nucleotides, sugars, steroids, etc. Cells contain receptors in their plasma membranes (usually transmembrane proteins), which bind signal molecules specifically, thus initiating a reaction cascade in the cell cytoplasm. This reaction elicits a change in cell behavior, e. g., a change in metabolic activities, in gene expression, in cell structure or motion. Signal-emitting cells are classified by the range over which their signals travel: paracrine cells secrete signal compounds short-range into the ambient interstitial liquid, whereas endocrine cells secrete hormones which are transported by blood over a long-range. Neurons show the highest organization with long axons, which, through synapses with other neurons, can contact far-away target cells and influence their behavior through secretion of neurotransmitters and their receptors. The transport of electrical signals (action potential), of ions and small metabolites among cells proceeds via gap junctions (cell-cell channels) which are made up of membrane-embedded protein complexes, the connexons. They make it possible for neighboring cells to participate collectively in signaling. Each cell type is equipped with a specific kit of receptors, resulting in a cell-specific answer towards extracellular signals.

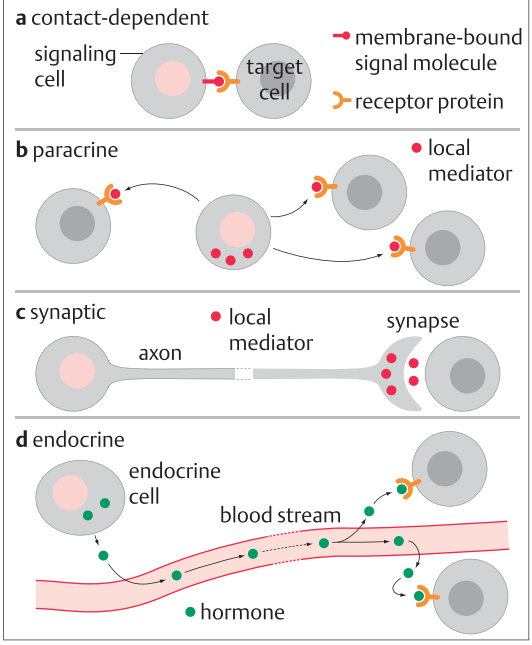
Types of cells



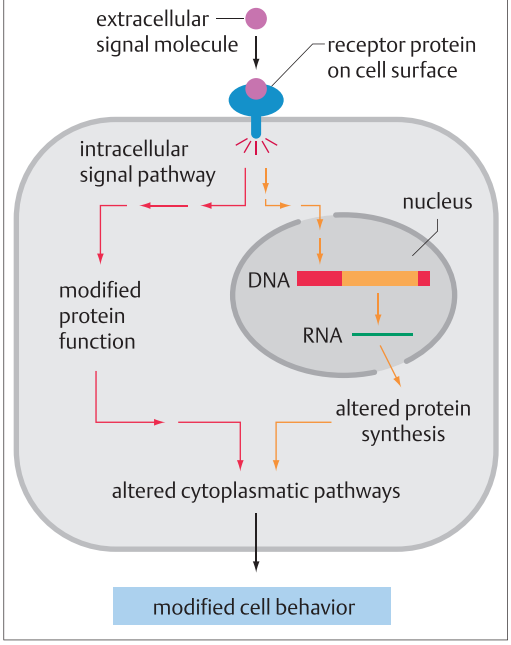
Transport through a lipid double layer



Intercellular communication



Intracellular signal transduction



Stem cells

General. Stem cells have the capacity to divide continuously, if held in culture, and to develop into various kinds of specialized cells. Embryonic stem cells appear in the fertilized egg during an early stage of development, and adult stem cells occur in most tissues of adult animals or humans. Stem cells are an important tool in fundamental research, as they may teach us the molecular events that occur during development. They also may have great therapeutic potential in treating diseases related to tissues or organs (cell therapy). A breakthrough occurred with the discovery that fully differentiated cells can be turned into multipotent stem cells (→306) if appropriate transcription factors are used. These stem cells can be induced to become a wide variety of differentiated cells (iPS = induced pluripotent stem cells).

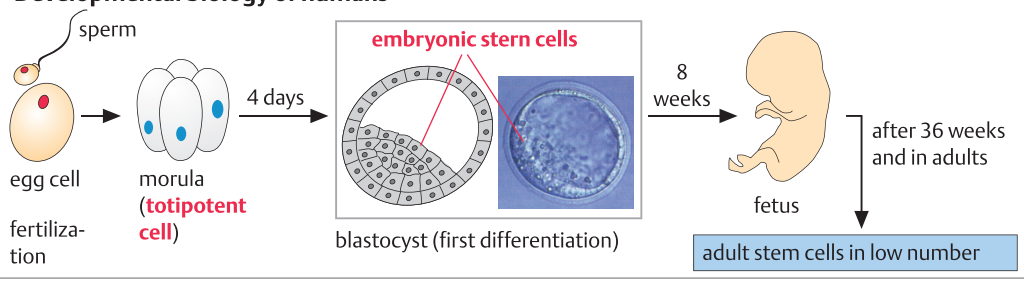
Embryonic stem cells (ECS). All cells developing from a fertilized egg cell have in their early (morula) stage the capacity to differentiate into any kind of specialized cell (they are totipotent). Homozygous twins are the natural consequence of two totipotent cells separating from one morula. By ca. 4 d after fertilization, the morula has developed into a blastocyst, whose inner cells are multi- or pluripotent – they are still able to form a wide range of different cell types (→266) – and whose outer cells have already started to differentiate. Upon further cell division, the inner cells form a large reservoir of multipotent ECS that are able to differentiate into a wide range of specialized cells, e. g., bone marrow, nerve, or heart muscle cells. In the human embryo, this development is complete after ca. 8 weeks; most ECS by now have differentiated. As a consequence, human ECS may be isolated 1) from human blastocysts that have been generated by *in-vitro* fertilization (IVF) of infertile couples but have not been implanted; 2) from fetal tissue after miscarriages or abortions; 3) by transfer of the nucleus of any diploid human cell into an enucleated human egg cell and cultivation of this cell to the blastocyst stage (see also "cloned animals") (→266).

Adult stem cells. The bone marrow of children, and even of adults, contains multipotent stem cells. In low numbers, they reach the circulatory system and differentiate into various types of blood cells. More recently, stem cells have also been found in many other tissues of the adult human, e. g., neuronal stem cells in dissected brain samples. Based on animal experiments, it is now believed that adult stem cells, if transplanted into a different type of tissue, may adapt their differentiation to the host tissue, thus behaving like multipotent cells. A great advantage of adult stem cells is their immunocompatibility, provided the donor and acceptor are the same person. However, they are much harder to isolate than ESC. In addition, they possess the inborn or acquired genetic defects of the donor. It thus seems at present that ECS have a far wider application potential than adult stem cells. A breakthrough occurred in 2007 with the discovery that fully differentiated cells can be reprogrammed into iPS, induced multipotent stem cells.

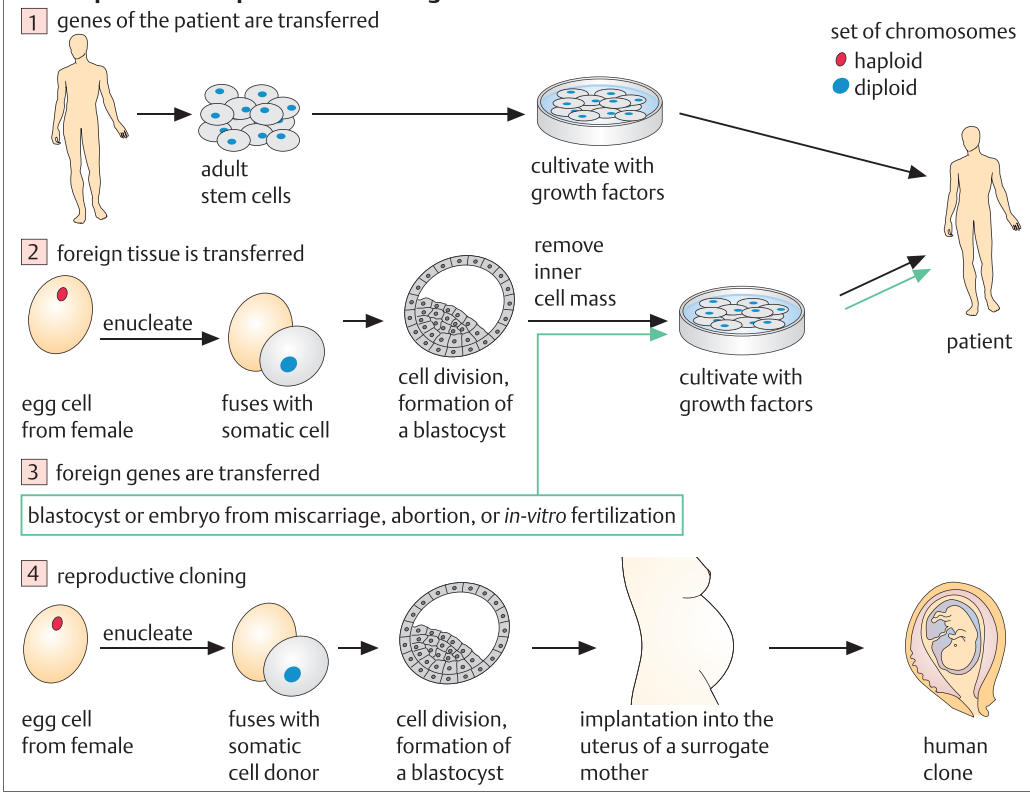
Applications. ECS permit fundamental research on the molecular basis of cell differentiation to be carried out. In addition, they are a valuable tool for studying pathological situations, e. g., birth defects or tumors. They also could serve to develop a wide range of human cell lines most useful for testing drug efficacy and safety on a molecular level. Finally, their use opens the possibility of curing diseases through cell therapy. For example, the transplantation of pancreatic island cells, obtained from cultivation of stem cells in pancreatic tissue culture, could permanently cure children suffering from type I diabetes. It must be said, however, that many technical questions are still unresolved, e. g., reliable differentiation of ECS *in vitro* and their immunocompatibility with the host.

Ethical concerns. Whether human life has started already at the multicellular stage of a fertilized egg such as the morula or the blastocyst, which is subject to legal protection, is a controversial issue and subject to ethical debates. In the USA and most industrialized nations, a limited acceptance of embryonal stem cell research in therapeutic cloning has been reached, which is paralleled by emphasizing research on adult stem cell cloning and therapy. With the advent of iPS cell technology, these concerns have been toned down; the origin of iPS cells are somatic cells from the donor, and after dedifferentiation and redifferentiation, he or she is also the recipient. However, concerns remain as to the safety of this new form of therapy.

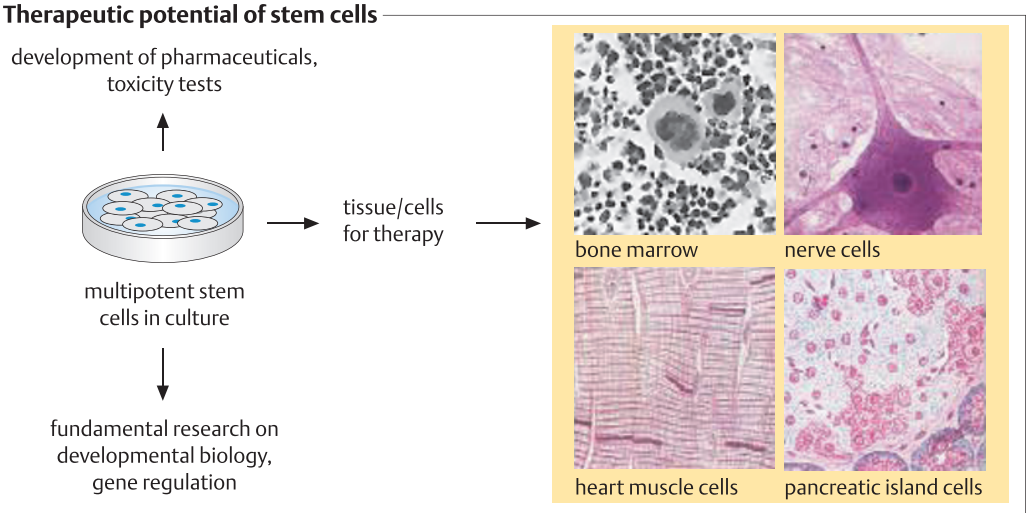
Developmental biology of humans



Therapeutic and reproductive cloning



Therapeutic potential of stem cells



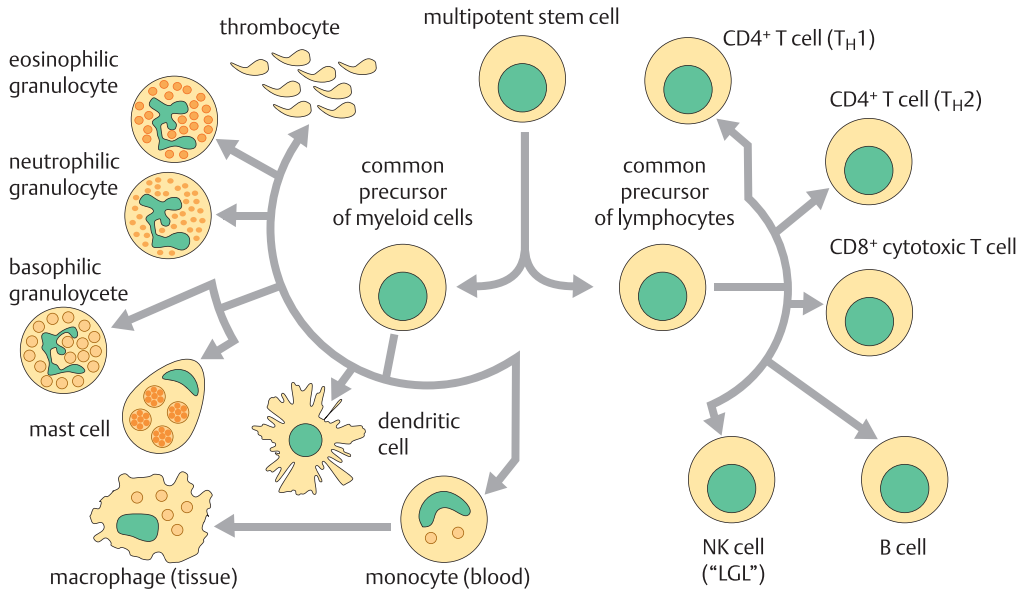
Blood cells and immune system

General. The immune system protects higher organisms from infections and provides immunity against many pathogens. It consists of specialized cells (cellular immune response) and messenger chemicals (humoral immune system) that communicate with them. Cytotoxic cells of the immune system destroy pathogens that have invaded the body, and also native cells of an organism that have been irreversibly damaged (apoptosis). They also participate in the immune defense against transplanted organs (→272). To comply with changing environmental conditions, the immune system displays high plasticity, which is genetically determined. Misguidance may lead to a wide range of diseases such as insufficient immune response, allergies, autoimmune diseases, and malignant degeneration. The immune system is regulated by many messenger proteins (cytokines, growth factors). Many of them can be prepared as recombinant proteins and are being evaluated for therapeutic use.

Cell types. All types of blood cells with functions ranging from oxygen transport to antibody formation are generated from a common stem cell precursor (→78): the multipotent hematopoetic stem cells, located in the bone marrow. They differentiate there into myeloic and lymphatic stem cells. The former give rise to red blood cells (erythrocytes), which stay in the blood vessels and transport oxygen, bound to their hemoglobin, as well as to granulocytes, macrophages, and other cell types. The lymphatic stem cells, however, develop into the lymphocytes that emigrate into the blood and lymph system. A healthy adult possesses ca. 10^{12} of these "naive" lymphocytes (meaning those which have not yet come into contact with antigens). Once a lymphocyte has been activated by an antigen (and some other signals), it forms, by clonal expansion, a large number of antigen-specific daughter cells. Lymphocytes differentiate further into B cells and T lymphocytes. Once B cells have matured in bone marrow, in lymph nodes or in the spleen, they form antibodies upon contact with an antigen (humoral immune response). In contrast, T cells mature in the thymus, where they differentiate upon contact with molecules of the major histocompatibility complex (MHC), a protein complex of the cell membrane which is exposed on the cell surface. The MHC-T-cell complex forms specific surface structures, which are distinguished by their function. T-cells are the main carriers of the cellular immune response. They secrete different cytokines. As an example, T helper cells may secrete various interleukins, thus activating, expanding and differentiating B-cells. CD4 is a typical glycoprotein marker on the surface of T helper cells. In contrast, cytotoxic T lymphocytes carry the glycoprotein CD8 on their surface. They can lyse virus-infected cells and secrete, among other substances, the cytokines interferon-γ and lymphotoxin-α.

Immune response and cytokines. The immune response to infections differs, depending on whether viruses, bacteria, or parasites are the pathogen. Extracellular pathogens or their toxins are first tagged by antibodies, triggering a cascade that results in their endocytosis and degradation by macrophages. Intracellular pathogens, such as mycobacteria or viruses, are destroyed by a different mechanism (similar to the elimination of transformed cells): as soon as they have infected one of the omnipresent macrophages, it will expose lysed fragments of the pathogen/cell on its surface, initiating a complex cascade that results in the destruction of infected cells by cytotoxic T lymphocytes. Autoimmune diseases follow a similar mechanism. For example, in type I diabetes, proteins of the β cells of the pancreas have been modified and are consequently misinterpreted as foreign proteins, resulting in their destruction by CD8 lymphocytes. The coordination of the immune response is largely effected by the cytokines and their receptors on the surfaces of cells of the immune system. Regulation of the immune response is highly complex. Cell-specific growth regulators and their receptors determine in a highly specific manner which cells of the immune system must be synthesized at a given time. The advent of genetic engineering has led to the possibility of producing cytokines and growth factors as recombinant proteins, initiating a novel area of medical research and, in some cases such as the interferons, new possibilities for medical therapy. Antibodies and stem cells are alternative technologies which are being evolved for immune therapy.

Blood formation and immune response: cell types



Cytokines and growth factors (selection)

general function	type	from cell type	tragets and effects
activation of lymphocytes	interleukin-2	T_H1, (CTL)	promotes growth of T-cells
	interferon-γ	T_H1, CTL	activates macrophages
	interleukin-4	T_H2	activates B cells, promotes growth of B- and T-cells
	interleukin-3	T_H1, T_H2, (CTL)	stimulates growth of hematopoetic precursor cells
local inflammation	interleukin-9	T-cells	increases mast cell activity
	interferon-α	leukocytes, fibroblasts	increases the expression of MHC class-I molecules
	TNF-α	macrophages, NK cells	causes local inflammation reactions
systemic and bone-marrow-specific effects	interleukin-1α, interleukin-1β	various cell types	causes fever, promotes growth of hematopoetic precursor cells
	interleukin-6	T_H2, macrophages	liberates acute-phase proteins
	erythropoietin	kidney	stimulates growth of erythroblasts
	granulocyte-macrophage colony-stimulating factor (GM-CSF)	T_H1, (T_H2), (CTL)	increases generation of granulocytes, macrophages, and dendritic cells

Therapeutic potential of cytokines and growth factors

infections	tumor treatment
shock	autoimmune diseases
defects in the immune system	allergic reactions
defects in cell growth	transplantation medicine

Antibodies

General. Antibodies are specific defense proteins circulating in the blood and lymph of vertebrate organisms. They are formed upon contact of B-lymphocytes with immunogenic antigens and bind with high affinity to such antigens. Most foreign proteins, polysaccharides, and lipopolysaccharides can act as antigens, e. g., macromolecules that constitute the cell surface of viruses, microorganisms, and parasites. Toxic proteins (toxins) may also lead to antibody formation. Even low-molecular-weight compounds may give rise to antibody formation if they are presented on the surface of strong immunogenic structures ("haptens"). In autoimmune diseases, the organism's own proteins have become "foreign" and have developed antigenic properties. Antibodies have long been used as vaccines for the treatment of infections and toxins (e. g., snake bites) (passive immunization) (→248). They are also of great value as reporter groups (→84) in immunoanalysis (→260). They are sometimes used for the purification of recombinant proteins, e. g., factor VIII, by immunochromatography (→186).

Structure. Antibodies belong to the immunoglobulins. In man, they are classified into 5 groups (IgG, IgM, IgA, IgE, and IgD), which play various roles in immunodefense. IgG, which predominates in serum, is a glycosylated heterodimer composed of two identical light (L) and two heavy (H) chains, which are linked by cysteine bridges. Structures having constant (C_H, C_L) and variable sequences (V_H, V_L) domains can be distinguished in the heavy and light chains. The F_c-region of the antibody binds to a receptor, and the F_{ab}-region binds to the antigen. This region of the antibody is hypervariable: the 6 complementarity-determining regions (CDRs) consist of ca. 20 amino acids each; thus, each CDR allows for $20^{6\times 20}$ permutations.

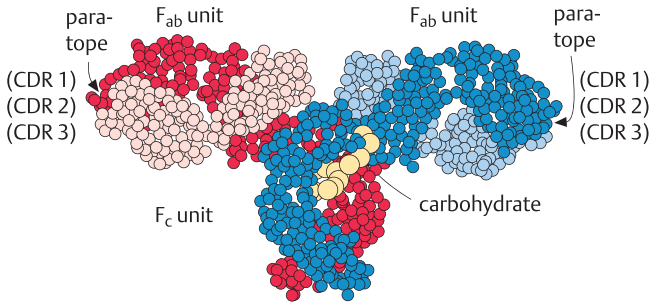
Biosynthesis. Antibodies are synthesized by B lymphocytes, which have nearly 1,000 sets of gene segments available. The gene segments are combined by random recombination ("gene shuffling") (→198) to code for the variable region of the immunoglobulins. In addition, during the expansion of B-cell clones, mutations occur in the genes that are responsible for the variable regions. Thus, a relatively small genotype coding for antibodies is turned into a huge phenotypic diversity.

Preparation. Polyclonal antibodies are mixtures of different antibodies that are directed towards different epitopes of the same antigen. They are obtained by immunization of animals such as rabbits, sheep, goats, and horses. Through repeated immunization, in intervals of several weeks, and extraction of blood, similar lots of antibodies can be obtained repeatedly from the same animal (horse, cattle, sheep). Purification is done by precipitation and chromatographic procedures. In manufacturing highly purified antibodies, affinity chromatography based on protein A may be used. Protein A (M_R 42 kDa) is obtained from *Staphylococcus aureus*. It binds with high specificity and affinity to the F_c-region of IgG. Purified IgG solutions are portioned in a sterile manner and lyophilized in the absence of air. If stored under refrigeration, antibodies are stable for several years. Industrial production is done under GMP conditions.

Risks. For human therapy, antibodies are administered parenterally, since they are not stable to gastrointestinal passage. Antibodies obtained from animals are recognized by the human immune system as foreign and thus can give rise to an immune defense, especially after repeated injections. One solution to this problem is to shift among antibodies obtained from different animal species. Alternatively, antibodies can be obtained from blood donors. Although donated blood stored in blood banks is thoroughly scrutinized before use, a risk of viral contamination such as hepatitis or AIDS does exist. Recently, monoclonal (→242) or recombinant antibodies (→244) or antibody fragments can be obtained through cell technology. Their advantage is that a well defined antibody can be produced in a bioreactor in unlimited quantities. In order to cope with immunogenicity, hybrid structures with the epitopes and sugar chains of human antibodies have been developed ("humanized antibodies"). Human antibodies can now also be manufactured at a large scale using cell technologies.

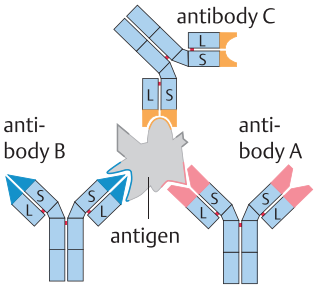
Applications. Antibodies are mainly used for diagnosis and therapy of human diseases and as analytical tools in molecular and cell biology (→246). More recently, they are also being applied in food and environmental analysis.

IgG antibody



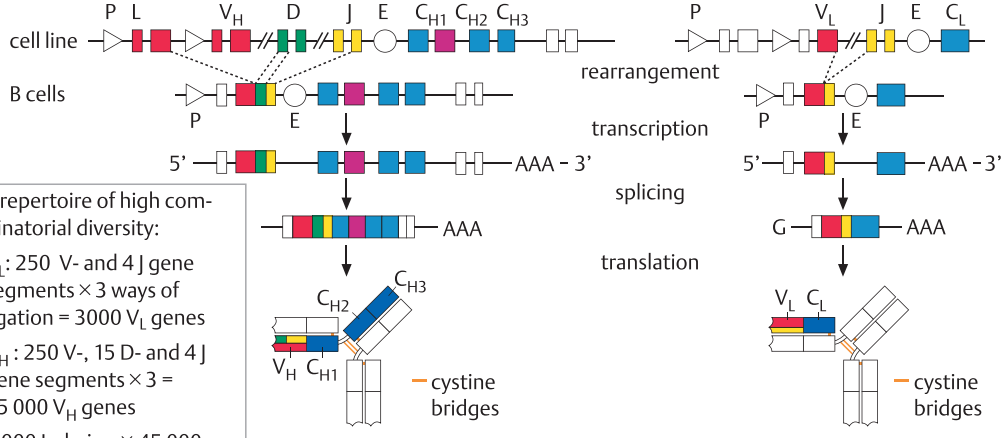
● ● heavy chains (H)
○ ○ light chains (L)
○ carbohydrate

CDR 1 – 3 = complementarity-determining regions, where antigen binding takes place



various paratopes of the antibody
various epitopes of the antigen

Biosynthesis



a repertoire of high combinatorial diversity:

V_L: 250 V- and 4 J gene segments × 3 ways of ligation = 3000 V_L genes

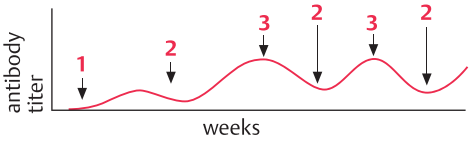
V_H: 250 V-, 15 D- and 4 J gene segments × 3 = 45 000 V_H genes

3 000 L chains × 45 000

V_H chains = $1{,}35 \times 10^8$ antibody variations, further increased by somatic mutations during affinity maturation

P = promoter
L = leader sequence
V_H, V_L = variable domain of the heavy or light chain
D = diversity segment
J = "joining" element
E = enhancer
C_{H1-3}, C_L = constant domains of the heavy or light chains

Polyclonal antibodies



1 antigen is injected
2 injection is repeated
3 blood containing antibody is collected

Purification by chromatography

several L of blood
separate erythrocytes by centrifugation, obtain plasma

IgG fraction
precipitate with ethanol or ammonium sulfate

purificcation
chromatographic procedures, affinity chromatography using protein A

Commercially available antibodies

some commercially produced polyclonal antibodies for passive immunization

antibody	obtained from
tetanus antitoxin	horse serum
snake venom antisera	horse serum
measles virus immunoglobulin	human serum
immunoglobulin G	human serum

Reporter groups

General. Reporter molecules play an important role in both basic and applied research. They are used 1) in cytochemical analysis of cells and histochemical analysis of tissues, 2) for sorting cells in a cell sorter, 3) for visualization of binding events, e. g., of antibodies, receptors, or DNA (→60, 260), and 4) in many procedures used in genetic engineering (e. g., for cloning or the investigation of promoters). Frequently used reporter atoms and molecules include radioactive isotopes, fluorophores, and enzymes. For linking reporters to proteins or DNA, the biotin-streptavidin and digoxigenin systems are often used. Frequently used reporters in genetic engineering are the genes coding for β-galactosidase, luciferase, and GFP (green fluorescent protein).

Radioactive markers. In radioimmunoassays (RIA), reactants are labeled with ^{131}I or ^{35}S, and radioactivity is determined in a scintillation counter. In molecular genetics experimentation, ^{32}P-labeled phosphate is often introduced into DNA or RNA, using DNA polymerase I, the Klenow fragment, or RNA polymerase. Radioactively labeled DNA or RNA can then be used to detect hybridization events by autoradiography or, somewhat faster, in a commercial phosphoimager (→48). Radioactive methods are still being used in many laboratories, in spite of strict radiation-safety regulations and competition from faster protocols, because they are highly sensitive.

Fluorophores. Fluorophores such as fluoresceine or rhodamine are highly sensitive reporter molecules, down to the picomolar range. Fluorophore-labeled antibodies are used in histochemical and cytochemical investigations. In combination with a cell sorter (FACS = fluorescence-activated cell sorter), they permit labeled cells to be separated from unlabeled cells at speeds of 1,000 cells min^{-1} or more. This procedure is used in cell biology, e. g., for separating B cells from T cells after labeling specific antigens displayed on their surface, and for separation of mixed cultures of bacteria after labeling taxon-specific sequences of their 16S rRNA by hybridization (FISH = fluorescence in-situ hybridization) (→74). Fluorescence reporters for DNA are SYBR-Green and ethidium bromide (→50).

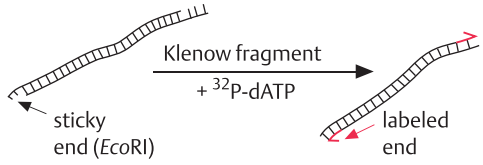
Enzymes. Compared to the reporter molecules mentioned so far, enzymes have the advantage of further enhancing an assay's sensitivity by signal amplification. This property is especially useful for quantitative analytical determinations. Reporter enzymes that are used often are alkaline phosphatase and horseradish peroxidase (→256). Their reactions can be assayed in biosensors, by electrochemistry, or photometrically (→258). If fluorometric or chemiluminometric enzyme assay protocols are used, the sensitivity may reach pico- or even attomolar detection limits. Thus, a very sensitive protocol is the use of horseradish peroxidase as a reporter enzyme in combination with luminol, a phtalic acid derivative which shows strong luminescence upon oxidation.

Digoxigenin and biotin-streptavidin. These chemicals are often used to couple a reporter molecule to the biomolecule to be analyzed. Digoxigenin (M_R = 390.52) is a steroid that can be coupled to a nucleotide via its hydroxy group without interfering with hybridization events. It binds with high affinity (10^{-9} M) to a specific antibody that can be labeled with various reporter molecules. A similar system is biotin-streptavidin (affinity 10^{-10} M).

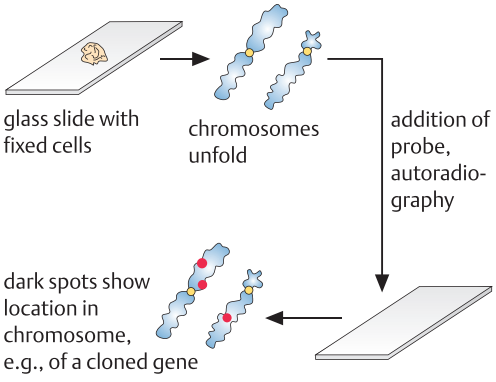
Genetic markers. Introducing genes that code for reporter proteins enables cloning events to be rapidly analyzed. "Blue-white" screening (→62), for instance, is based on a plasmid-borne DNA sequence coding for a fragment of β-galactosidase. If foreign DNA is inserted within this fragment, it can no longer complement the chromosomal DNA that codes for the remaining part of this enzyme, and externally added chromogenic substrate X-gal (5-bromo-4--chloro-3-indolyl-β-D-galactopyranoside) can no longer be hydrolyzed to a blue dye. Another very useful genetic marker is GFP (green fluorescent protein), whose gene was cloned from a jellyfish. Without any additional substrates, GFP shows strong fluorescence. The GFP gene can be used as a reporter to demonstrate promoter function, gene expression, or gene regulation. Luciferase genes cloned from fireflies or photobacteria are used for the same purpose, but visualization of the expressed reporter proteins requires the addition of an enzyme substrate (luciferin or decanal, respectively).

Radioactive labeling

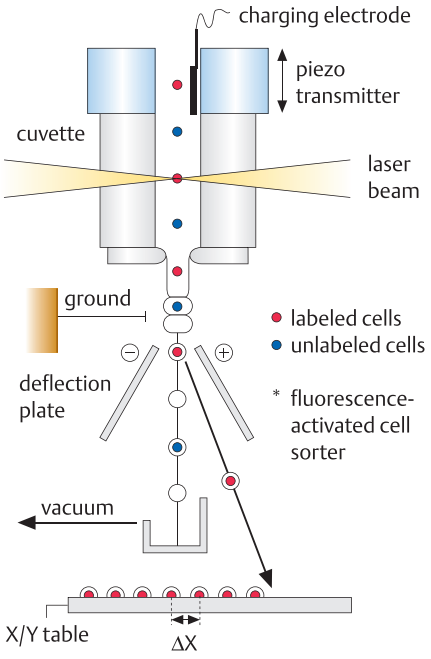
by filling up with ^{32}P-dATP



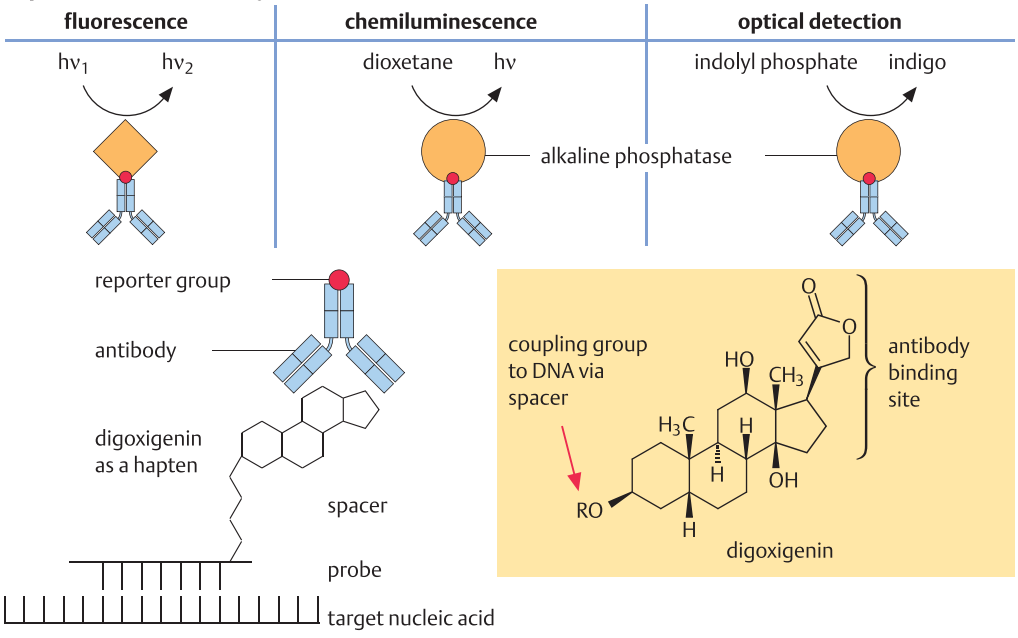
hybridization with radioactive DNA probe



Cell sorter (FACS*)



Optical tests for DNA hybridization



Frequently used reporter genes

gene	externally added substrate	detection
lacZ': gene fragment of β-galactosidase	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal)	visual (blue-white)
lux: luciferase from firefly or photobacteria	luciferin or decanal	luminescence
GFP: green fluorescent protein from jellyfish	–	fluorescence

Solid state fermentation (SSF)

General. Solid state fermentation, often defined as "fermentation without the presence of free water," is a traditional fermentation procedure still widely used for the preparation of fermented food. Important examples are the production of soy sauce, miso or tempeh (→114). In the early days of biotechnology, it was also widely used for the manufacture of fungal products such as citric acid (→146) or food enzymes (→172), but has been largely replaced in industry today by submerged fermentation. Recently, SSF has been rediscovered and is being explored for the manufacture of bioethanol and similar compounds in the framework of "biorefineries." (→138)

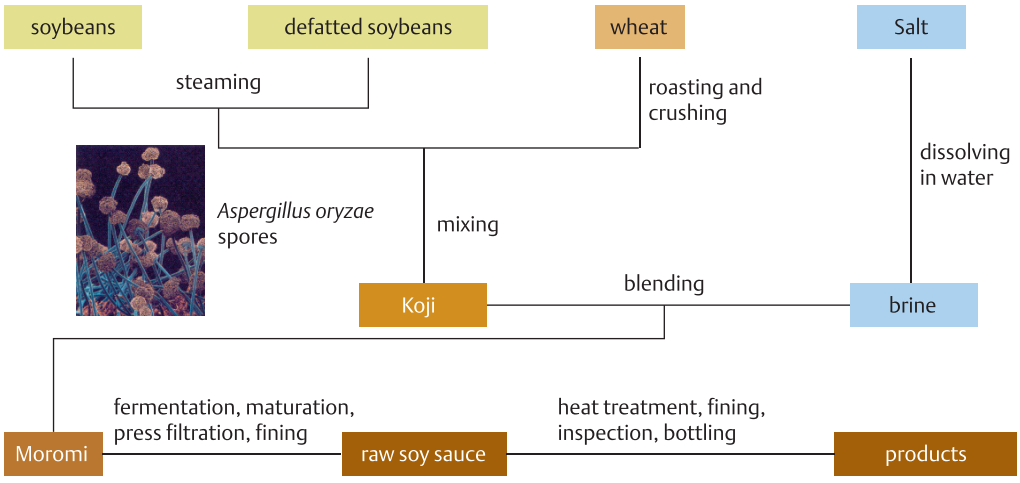
Raw materials. Generally speaking, SSF procedures are mostly used to add value to waste materials such as wheat bran, soy or coffee bean residues. Before inoculating with starter cultures (→114) or waiting for the growth of airborne microorganisms, the raw materials are sometimes pretreated in order to enhance the concentration of digestible sugar.

Microorganisms. SSF is a traditional domain for fungi such as *Aspergillus* strains (→18), because hyphal growth, as with other fungi, allows for an excellent penetration of substrates and thus for good product yields. However, yeasts (ethanol), Lactobacilli (yogurts) (→116) or Bacilli (Natto) are also being employed.

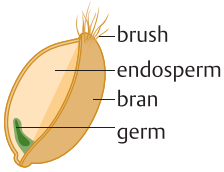
Technology. In the preparation of traditional food products such as, e. g., Sauerkraut (→116), Tempeh or cocoa, traditional crafts are still being used. For industrial manufacturing of products, considerable work on engineering and process control has been done. As the key limitations of SSF are mass transfer (→94), removal of reaction heat, analytical controls and process modeling, much work has been directed towards these issues. Mass transfer usually relates to substrate access and, in aerobic microorganisms, to oxygen. Both needs can be satisfied by putting the substrate-microorganism mixture on gently agitated percolated trays. Removal of reaction heat is a significant challenge and has been addressed by air blowers, internal heat transfer plates or water cooling jackets. On a solid substrate, water activity must be controlled as it is often critical both to microbial growth and product formation. Finally, process analysis (→96) in a solid system is not straightforward, and methods such as microcalorimetry, Fourier-transform or near infrared analysis (FTIR, NIR), aroma-sensing devices such as GC or GC-MS, water content quantification by microwave analysis, or image analysis have been employed. Many of these procedures are under study, and experimental reactor designs have been realized, for example, in a horizontal paddle mixer established by Wageningen University, in the Zymotis reactor developed by the ORSTOM Center of Montpellier or by the Platotex bioreactor engineered at the CNRS in Gif-sur-Yvette. All these designs are an attempt to develop SSF into a highly reproducible procedure, and advances have been made in process modeling, but, if compared to submerged fermentation, few processes have reached maturity. In summary, SSF has the advantages of low-cost media, a rather inexpensive bioreactor (e. g., tray) design, few operational problems and simple downstream processing (→108) (usually water extraction), but there are difficulties in scale-up and process control. SSF is used for the production of food enzyme mixtures in a koji process (e. g., starch-degrading enzymes, pectinases), and for the manufacture of some important traditional food additives such as soy sauce.

Soy sauce manufacture (Japanese style). (→114) Traditional raw materials for soy sauce preparation are a hot-water extract of soybeans and an equal amount of roasted wheat. This is mixed with a starter culture of *Aspergillus oryzae* (sometimes also *A. sojae* or *A. tamari*) resulting in a "koji" where Bacilli, yeast and Lactobacilli are also present. Once a microbial community has developed, the koji is mixed with coarse salt (SSF) or with brine (wet fermentation) and brewed for as long as several months in "moromi" fermentation tanks. During moromi brewing, *A. oryzae* breaks down the grain and soy proteins into free amino acids and oligopeptides, and the wheat starch into simple sugars. Free amino groups and sugar aldehyde groups interact in the so-called Maillard reaction, leading to a dark-brown color. Lactic acid bacteria ferment the sugars into lactic acid, and *Saccharomyces* produces some ethanol, which through aging and secondary fermentation creates numerous flavor compounds typical of soy sauce. This complex process has been thoroughly analyzed. The metabolic pathways of the standardized microorganisms involved (starter cultures) are well understood, and each single process step has been optimized. As a result, soy sauce can be brewed in a SSF process at a reproducibility similar to beer brewing by submerged fermentation. There are many varieties of soy sauce flavors throughout Asia.

Manufacture of soy sauce by solid-state fermentation



Frequently used substrates



wheat grain

wheat bran

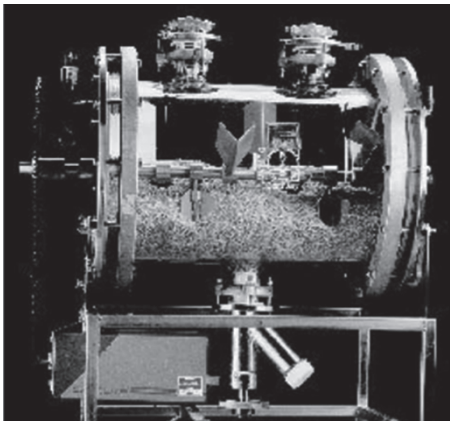
wheat, rye or rice bran
cassava
sugar beet molasses
sugar cane molasses
sweet potato residues
coffee pulp and husk
soybean hulls
apple pomace

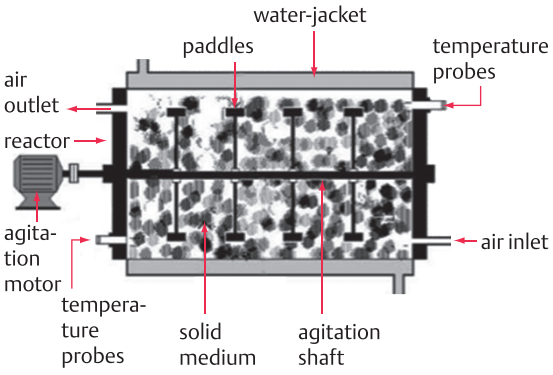
Frequently used microorganisms

microorganism	product
Aspergillus oryzae	miso, soy sauce
Aspergillus niger	citric acid
Rhizopus oligosporus	tempeh
Acetobacter pasteurianus	wheat-based vinegar
Trichoderma sp.	composting, silage
mixed cultures	cocoa bean fermentation

Fermentation of cocoa beans

Horizontal paddle mixer for solid-state fermentations





Wageningen University, The Netherlands

Growing microorganisms

General. Microorganisms are cultivated either on solid nutrients (surface cultivation) or in liquid culture (submerged cultivation). In laboratory experiments, agar plates or shake flasks prevail. Under industrial conditions, bioreactors are the cultivation vessels of choice. The composition of the culture medium is of key importance for product formation. In most cases, contamination by undesired microorganisms is excluded by using sterile cultivation conditions.

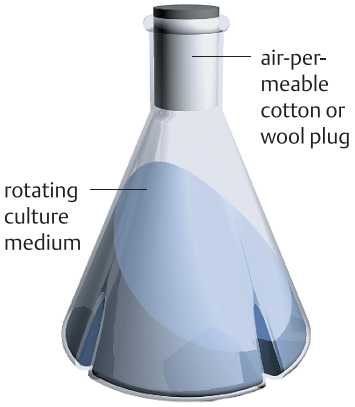
Shake flasks. The standard vessels are Erlenmeyer flasks with baffles (filled with ca. 50–500 mL liquid), containing a sterile liquid nutrient solution. Oxygen saturation is assured by shaking the flasks on a thermostatted reciprocal or gyrating shaker. When anaerobic bacteria are to be grown, oxygen is removed from the nutrient medium by boiling and degassing, followed by addition of thiogycolate. Further handling is done under an O_2-free sterile hood.

Bioreactors (fermenters) are closed reactors with a capacity of 1 L to > 500 m^3. In the standard, stirred bioreactor, mass transfer and air distribution are accomplished with a stirrer. Bioreactors can be operated as batch cultures, as batch cultures with subsequent addition of substrates (fed-batch culture), or as continuous cultures. In industrial practice, batch and fed-batch cultures are preferred, whereas in fundamental studies, continuous culture is of great importance because it allows cells to be kept at a constant specific growth rate for many days or even weeks. In many microbial fermentations, product formation begins only in the late logarithmic phase of culture growth. If at this point more nutrients are added in fed-batch mode, the production phase of the fermentation can be prolonged, and the yield of the end product is increased. Another reason to use fed-batch fermentations is to prevent substrate inhibition: often, microorganisms produce less product in the presence of high glucose concentrations (catabolite repression).

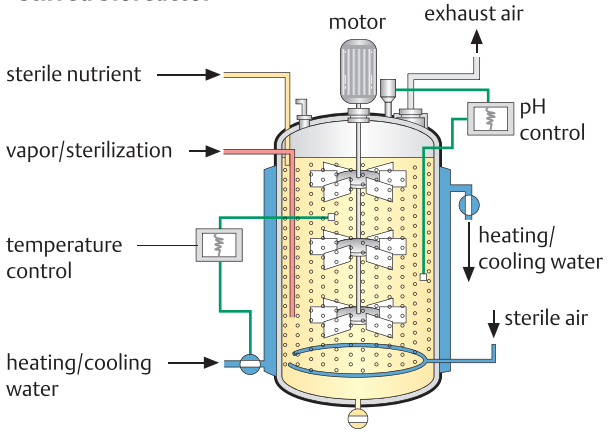
Medium optimization. Most microorganisms that are used in biotechnology are heterotrophic and grow aerobically. They require organic compounds as a source of carbon and energy, in addition to inorganic or organic nitrogen, salts, and trace elements. The nutrient medium is usually optimized in shake flasks, using product yield and substrate cost and availability as the major parameters (in some fermentations, such as for ethanol or citric acid production, the cost of the C source may exceed 50 % of the production costs). For cost reasons, most industrially used nutrient media are composed of components that are not very well defined, such as corn starch hydrolysate, molasses, or soy meal (complex media), whereas in the research laboratory, defined media components such as glucose or mixtures of amino acids are preferred.

Sterilization. Autoclaves are best for sterilizing nutrient media in the laboratory prior to inoculation. Autoclaving for 15 min at 121 °C is sufficient to kill even the spores of thermophilic microorganisms (test organism: *Bacillus stearothermophilus*). Heat-sensitive medium components such as glucose and vitamins are usually added via sterile filtration to the autoclaved medium after cooling. If a bioreactor exceeds ca. 10 L in volume, it is usually autoclaved in place with steam at 1.4–3 bar. The method is time-consuming (heating and cooling cycles of several hours) and, due to the long exposure to heat, may lead to changes in the medium composition. Thus, continuous sterilization is preferred in industry. In this procedure, the nutrient broth is exposed to steam at 140 °C for ca. 2–3 min (holding time). Using countercurrent heat exchangers, the formation of vapor condensate is prevented, and ca. 90 % of the introduced energy is recovered. Air must be purified by filters before it enters the bioreactor: 1 m^3 of air may contain up to 2000 colony-forming units (cfu), including up to 50 % fungal spores and 40 % Gram-negative bacteria. For a bioreactor of 100 m^3 working volume, operating at an aeration rate of 1 vvm (volume air/volume liquid · min), 6000 m^3 of sterile air is required per hour. In recent years, single-use bioreactors made from plastic have become popular for some applications. They are available in a volume range of 500 ml to 3,000 L and are supplied as sterile units. The advantages of single-use bioreactors are that a) more experiments can be run in the same time, as the time required for sterilization and cooling is eliminated, b) the risks from cross-contamination are greatly reduced, and c) the initial investments are less compared to steel reactors.

A shake flask



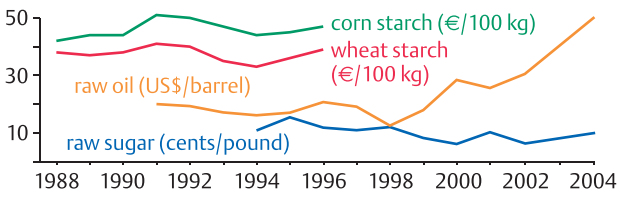
Stirred bioreactor



Nutrient components

component	origin	composition/remarks
complex carbon sources		
sugar beet molasses sugar cane molasses	sugar production	~ 48 % saccharose, ~ 33 % saccharose, ~ 22 % invert sugar,
corn steep liquor	corn milling	~ 1 – 3 % glucose, 11 – 13 % lactose
distillers´solubles	alcohol production	variable
starch and dextrins	corn starch hydrolysis	variable
sulfite suds	paper production	2 – 4 % hexoses and pentoses
whey	dairies	3 – 5 % lactose
hydrocarbons	petrochemical plants	aliphatic alkanes > C-5
defined carbon sources		
glucose		can be metabolized by most organisms
mannitol		good carbon source for Streptomycetes
methanol		can be metabolized by many bacteria and yeasts
complex nitrogen sources		
soymeal, peanut meal, wheat germ meal, corn steep liquor, whey powder, yeast extract		content of raw protein between 20 % and 6 %, contain vitamins and trace elements
defined nitrogen sources		
ammonia salts, nitrates, urea, amino acids		
vitamins, trace elements		
thiamin, riboflavin, pyridoxin, nicotinic acid, nicotinamide, pantothenic acid, cyanocobalamin, folic acid, biotin, α-lipoic acid, purines, pyrimidines, heme macro minerals (10^{-3} – 10^{-4} M): salts of K, Ca, Mg, Fe, S,P micro minerals (10^{-6} – 10^{-8} M): salts of Mn, Mo, Zn, Cu, Co, Ni, V, B, C, Na, Si		

Prices for some carbon sources



Prices are for the German market, end of 2013

BRENT crude, US-$/barrel	110
raw sugar, €/100 kg	40
corn starch, €/100 kg	57
wheat starch, €/100 kg	48

Growth kinetics and product formation

General. The rules governing the growth of microorganisms are well defined for single-cell organisms but not for mycelial organisms (Streptomyces, *fungi*). Several varieties of fermentation can be distinguished, depending on the kinetics of product formation.

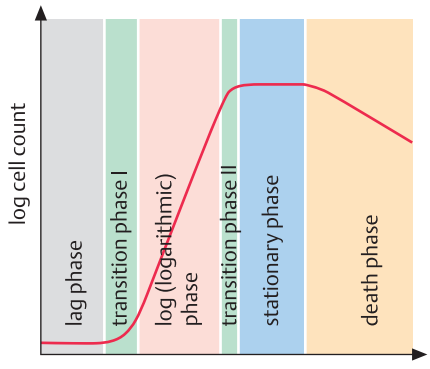
Growth kinetics of unicellular microorganisms. Most microorganisms and yeasts are unicellular. They propagate by cell division, and an increase in the number of cells can be monitored continuously by optical methods, e. g., turbidity. In a static culture, e. g., in a small shake flask or a batch reactor, a lag phase (when the formation of enzymes important for biosynthesis is induced) is followed, after a short transition phase, by a phase of logarithmic growth (log phase) having first-order kinetics. The following transition phase II is reached if one substrate becomes limiting or one product becomes inhibiting. This is followed by the stationary phase, where substrate limitation, excess population density, limited oxygen transfer, or the accumulation of toxic metabolites have terminated growth. A death phase characterized by decreasing cell number may ensue. To characterize a growth curve, important parameters are 1) the *lag phase* (dimension: [h]), which depends on the microorganism, the physiological conditions of the inoculation material, and the nutrient composition; 2) *the specific growth rate* μ (dimension: h^{-1}), which allows the rate of cell formation to be correlated with cell concentration during the exponential phase. When written as an equation $\mu = \mu_{max} \cdot S/(K_S+S)$ (the Monod equation), μ enables the experimental determination of the velocity of cell growth; 3) the *saturation constant* K_S of this equation relates to the substrate concentration (in mg L^{-1}), at which 50% of the maximum growth rate has been reached. From a formal point of view, K_S is equivalent to the K_M of enzyme kinetics, the Michaelis constant. 4) Growth rates are linked to the *generation* or *doubling time*. This parameter indicates in [h], how fast a bacterial culture doubles under exponential conditions. 5) The yield coefficient Y_S is a measure of biomass formation per consumed substrate. Various yield coefficients can be defined, because the formation of cell mass depends on both chemical (pO_2, C/N ratio, phosphate content) and physical parameters (e. g., temperature). If complex nutrient media are used, two log phases separated by an intermediate lag phase may be observed (diauxic growth). This is explained by the lag time required to induce new enzymes after the first carbon source is exhausted.

Growth kinetics of mycelium-forming microorganisms. Fungi, and also mycelium-forming prokaryotes such as Streptomycetes, grow not only by cell division but also by longitudinal growth of the mycelium. Determination of growth is usually done by weighing dried biomass and leads to complex kinetics.

Product formation in most fermentation processes is either coupled or not coupled to growth. Very few processes do not belong to one of these two basic types. In the traditional classification, growth-coupled products include e. g., cell mass formation (baker's yeast, SCP, algae) and the formation of primary metabolites such as ethanol, lactic acid, gluconic acid etc.. Product formation decoupled from growth occurs at the end of the logarithmic phase; the product does not arise from the primary, but rather from the secondary metabolism, e.g., antibiotics (→166) and extracellular enzymes (→200). If transgenic host organisms are used for production, the first step is usually to grow the host organism (*E. coli, S. cerevisiae*) to high cell densities which are in the range of 100 g/L culture medium (cell dry mass). Product formation is then coupled to growth if the gene coding for the desired product is cloned into a host chromosome and is constitutively expressed. Often, however, it is preferable to express the gene after the growth period is terminated. To this end, the gene is cloned behind an inducible promoter, e. g., the *lac* promoter which can be activated by allo-lactose (6-*O*-β--D-Galactopyranosyl-D-glucose) or isopropyl β-D-1-thiogalacto-pyranoside (IPTG) (→62). A *rhaPBAD* promoter is also often used; it is induced by the addition of L-rhamnose. If a $\lambda P_L P_R$ promoter is used, induction occurs after a temperature shift from 30° to 42° C – a procedure which is limited to experiments in laboratory scale since such temperature shifts are difficult to achieve in larger bioreactors.

Growth in a batch reactor

growth curve of microorganisms



	μ_{max} [h^{-1}]	t_D [h]
Escherichia coli, 35 °C	> 2	< 0,35
Saccharomyces cerevisiae, 35 °C	0.6	1.2
Aspergillus niger, 30 °C	0.2	3.5
Penicillium chrysogenum, 25 °C	0.12	5.7

growth kinetics in the exponential phase

$$\mu = \frac{1}{X} \cdot \frac{dX}{dt} \qquad \mu = \mu_{max} = \text{const.}$$

growth kinetics in transition phase II

$$\mu = \mu_{max} \cdot \frac{S}{K_s + S}$$

Monod equation (Monod correlation)
valid for substrate-limited growth

doubling time

$$t_D = \frac{\ln 2}{\mu}$$

generation time

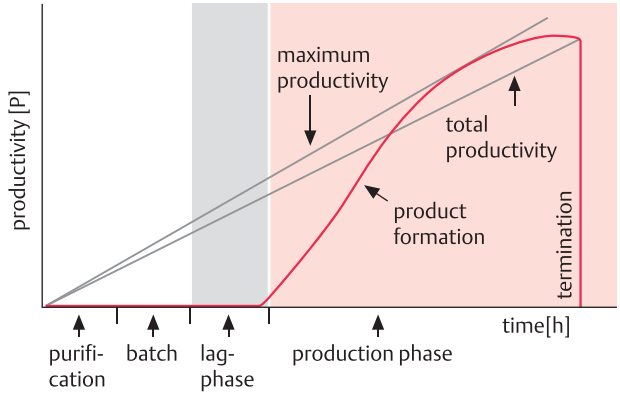
$$t_G = \frac{\ln 2}{\nu}$$

division rate

$$\nu = \frac{1}{N} \cdot \frac{dN}{dt}$$

ν = division rate [h^{-1}]
μ = specific growth rate [h^{-1}]
μ_{max} = max. specific growth rate [h^{-1}]
N = cell count [–]
X = biomass [g/L]
S = concentration of the limiting substrate [mol/L]
K_s = saturation constant, substrate-specific [mol/L] (Monod constant)
t = time [h]

Product formation



productivity

$$P = \frac{\text{product concentration}}{L \cdot \text{fermentation time}} \quad [\text{units}/(L \cdot h)]$$

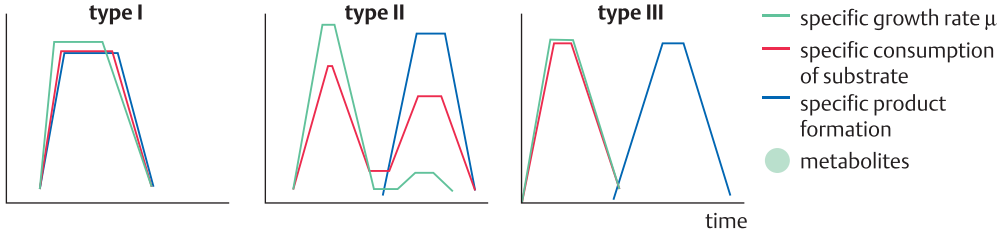
yield coefficients

$$Y_s = \frac{\text{biomass formation}}{\text{substrate consumption}} \quad [kg/kg]$$

$$Y_{O_2} = \frac{\text{biomass formation}}{\text{oxygen consumption}} \quad [kg/kg]$$

$$Y_{kJ} = \frac{\text{biomass formation}}{\text{heat formation}} \quad [kg/kJ]$$

Fermentation types



type I

substrate → product
or
substrate → A → B → product

examples: bakers' yeast SCP, ethanol

type II

A → B → C → primary metabolism
substrate → D → E → product

examples: citric acid, amino acids

type III

substrate → cell growth until end of log phase → product from secondary metabolism

examples: antibiotics, vitamins, enzymes

Fed-batch, continuous and high cell density fermentation

General. In fed-batch fermentations, the production phase is prolonged by feeding nutrient medium to the fermentation. This protocol is the preferred procedure in industry, followed by high cell density cultivation protocols for recombinant host organisms. Continuous fermentations are less practical, but of great fundamental importance, since they allow the laws governing microbial growth and metabolism to be studied.

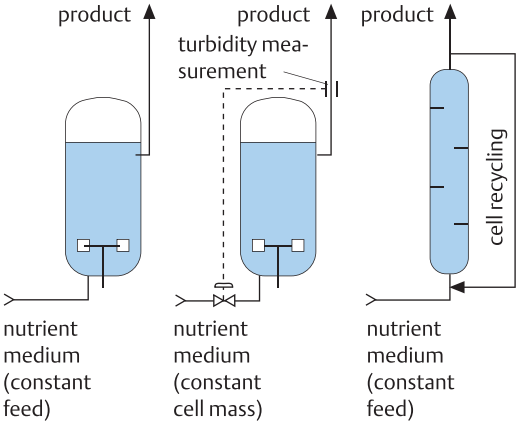
Fed-batch procedures have two important advantages: first, they increase the yields of the many secondary metabolites that are industrially produced (antibiotics, enzymes, polysaccharides, etc.) by providing fresh medium or intermediary building blocks at the end of the logarithmic phase just when secondary metabolism takes off. Second, substrate inhibition can be prevented by carefully limiting the glucose level in the medium. Glucose is the most widely used carbon and energy source in fermentation, but an excess represses product formation by catabolite repression, e. g., during production of antibiotics. Bakers' yeast production is another example of catabolite repression (→120): higher sugar concentrations lead to an increased specific growth rate μ; however, the biomass yield coefficient Y_S decreases strongly because an increasing amount of glucose is converted to ethanol (Crabtree effect). Thus, sugar is added in fed-batch mode to the fermentation broth. Similar requirements have sometimes been found for nitrogen and phosphorous sources.

Continuous fermentations. A batch reactor is usually considered a closed system, though strictly speaking there is gas exchange with the environment. During continuous fermentations, there is not only gas exchange with the environment, but the bioreactor is an open system to which sterile nutrient broth is continuously fed and from which culture medium is continuously removed. Three varieties of continuous fermentation modes are usually distinguished: two are chemostats, where nutrient levels are held constant, and turbidostats, where cell mass is held constant. The third is the plug-flow reactor, in which the culture medium flows without backmixing through a tubular reactor, while the cell mass is recovered at the reactor outlet and returned at the entrance of the reactor. In such a system, conditions along the direction of flow, e. g., medium composition, biomass concentration, and product concentration resemble the conditions that are obtained in a batch reactor over time. In a continuous fermentation under equilibrium conditions, cell loss from the effluent is balanced by the specific growth rate μ of the microorganism; and the substrate concentration S and the rate of product formation Q_X remain unchanged. Under such conditions, Q_X depends as a first approximation solely on the flow rate. Concepts for continuous fermentations are much harder to develop for secondary metabolites (type III fermentations) (→200), since cell growth and product formation are not directly coupled. Continuous fermentations are useful for 1) optimizing cell growth and product formation, and 2) analyzing the limiting nutrient components. In industrial practice, however, continuous fermentations are rarely used. Some exceptions are the aerobic and anaerobic treatment of wastewater (→286, 288), newer process variants for the production of beer, and the manufacture of human insulin using recombinant yeast strains (→222). In most industrial fermentation processes, however, it is held that 1) continuous processes, compared with batch-fed fermentations, show break-even economics only after 500–1,000 h of continuous operation – an extremely difficult condition to meet, from the point of view of sterile operation and anti-infection management; 2) it is difficult to keep the composition of nutrient media constant over such a long period of time; and 3) the genetic stability of recombinant host organisms cannot be assured over so many generations of cell division.

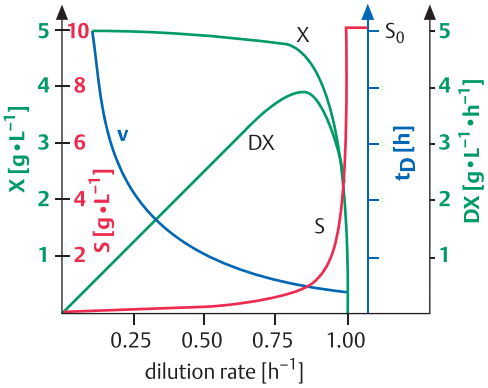
High cell density fermentation. This term is used if a concentration of at least 50 g/L cell dry mass is obtained in a bioreactor. Using appropriate feeding protocols, cell densities of *E. coli* up to 150 g/L cell dry mass can be obtained in bioreactors. The standard substrates are glucose or glycerol. For the efficient expression of the desired genes, a wide range of vectors with different induction strategies have been developed. High cell density fermentation protocols have also been developed for many yeasts, lactobacteria and other microorganisms.

Continuous fermentations

reactor types



dilution rate and yields



S = substrate concentration [g • L^{-1}]
X = biomass [g • L^{-1}]
t_D = doubling time [h]
DX = cell discharge [g • L^{-1} • h^{-1}]

biomass and substrate balances

biomass

$$\frac{dX}{dt} = \mu \cdot X - D \cdot X \qquad (1)$$

substrate

$$\frac{dS}{dt} = D\,(S_0 - S) - \frac{\mu \cdot X}{Y_S} \qquad (2)$$

for the stationary state it follows:

$$\frac{dX}{dt} = \frac{dS}{dt} = 0 \qquad (3)$$

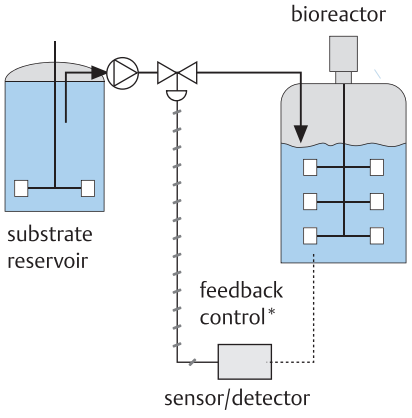
and it further follows that $\mu = D$

using (3) with (1) and (2):

$$X = Y_S\,(S_0 - S) \qquad D = \mu_{max}\,\frac{S}{K_s + S} \qquad S = \frac{D \cdot K_s}{\mu_{max} - D}$$

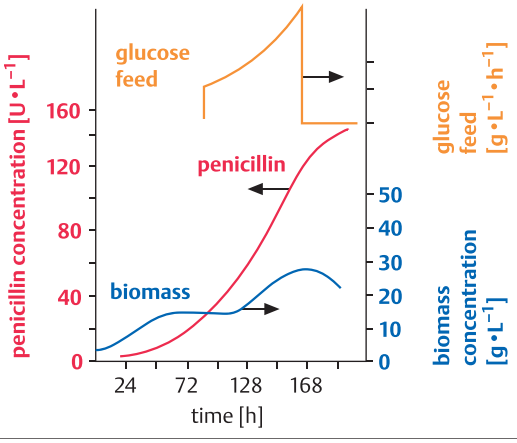
Y_S = yield coefficient [kg/kg]
K_S = Monod constant [g • L^{-1}]
S_0 = substrate concentration in feed [g • L^{-1}]
S = substrate concentration in bioreactor [g • L^{-1}]
D = dilution rate [h^{-1}]
X = biomass [g • L^{-1}]
μ = specific growth rate [h^{-1}]
μ_{max} = maximum specific growth rate [h^{-1}]

Fed-batch process



*e.g., by weight

penicillin production with glucose feed



Fermentation technology

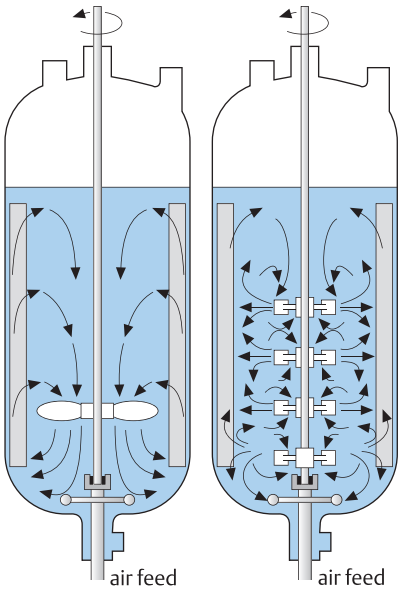
General. To produce biotechnological products at acceptable costs, bioprocess engineering, a discipline shaped by engineers, is as important as the biosciences developed by biologists and biochemists. Key objectives of industrial process engineering are the operational safety of a process and the minimization of both investment and process costs. Important aspects of these tasks are 1) optimized mass transfers; 2) technical solutions to keep temperature constant; and 3) optimization of aeration (for aerobic processes).

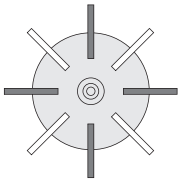
Mixing in a bioreactor is achieved by stirrers or pumps, resulting in a turbulent current. In the standard case of stirrers, turbulence in the immediate neighborhood of the stirrer is characterized by its Reynolds number Re. In aerobic processes, aeration also contributes to mixing. A factor in the numerical calculation of the Reynolds number is viscosity, which depends on the concentration of microorganisms, on their physical shape (e. g., mycelia in fungal fermentations), and on the type of product (e. g., xanthan). In an ideally mixed bioreactor, turbulence in the reaction zone is distributed in a homogenous manner. This target, however, is only approximated, because the sensitivity of the biological materials is usually limiting: stirrer speed, for example, is limited by the shear sensitivity of a mycelium. Various factors, such as the geometry of the stirrers, their form and number, the position of mechanical units such as baffles, the position of pumps (in unstirred reactors), and the configuration and positioning of aeration plates and air ejectors (in stirred and unstirred reactors). The power number Ne describes the energy requirement of stirred reactors and is correlated, in an ungassed state, with the Reynolds number. For industrial use, various stirrers have been developed, e. g., disc, turbine, MIG, and InterMIG impellers, which support good mixing and O_2 transfer. Mixing and O_2 transfer are measured by the volumetric transfer coefficient k_La.

Temperature control. For optimal results, fermentations are performed at constant temperature. After an initial heating as required for cell growth, fermentation reactors are usually cooled, because both microbial growth and stirrer movement produce heat, which must be removed. To calculate the heat produced in a process, the heat transfer number and the exchange area of the fermenter must be considered. It is usually sufficient to remove heat with a water cooling system that surrounds the bioreactor, but when yield coefficients are very high (low Y_{kJ} values), as in, e. g., alkane fermentations by yeasts, additional internal heat exchangers must be used. The reaction en-thalpy generated by a fermentation process can be roughly estimated from the degree of reduction of the substrate. Thus, the fermentation of alkanes generates more heat as compared to the fermentation of a partially oxidized substrate such as a sugar.

Aeration. The growth of aerobic cultures is usually limited by the oxygen content of the culture solution. To optimize the oxygen content, several biological and technical factors must be considered. For example, the optimal oxygen transfer in a bioreactor is correlated with the specific maximum oxygen uptake rate q_{O2}^{max} of the microorganism. In addition, oxygen is transported in a ternary-phase system comprising the gas and liquid phases and the microorganism. For oxygen transfer, several phase boundaries must be overcome: 1) from the gas bubble to the phase boundary surface; 2) through the phase boundary surface into the liquid; 3) through the liquid to the boundary surrounding the microorganism; and 4) into the cell. In single-cell organisms, phase boundary surface 2 is often limiting, whereas cell associations or mycelia often show limitations in 4. O_2 transfer also depends on a wide range of technical conditions such as reactor dimensions, hydrostatic pressure (filling height), stirrer performance, aeration system and rate, chemical and physicochemical parameters such as nutrient type, medium density and viscosity, temperature, surface pressure (antifoams), and finally on biological factors, e. g., the growth form of the microorganism. While the aeration rate can simply be given by volume air per volume reactor medium per minute (vvm), reactor and stirrer or pump geometry have significant effects on the oxygen transfer rate. A more meaningful parameter for the characterization of oxygen transfer in a bioreactor is thus the volumetric transfer coefficient k_La, which can be determined experimentally. Laboratory bioreactors reach 800 – 1200 h^{-1}, and production fermenters reach 500 – 800 h^{-1}.

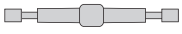
Mixing in a stirred reactor





disc impeller

turbine impeller



MIG impeller

InterMIG impeller

characterization of flow

Reynolds number Re

$$Re = \frac{\text{inertial force}}{\text{viscous force}} = \frac{d_R^2\ n\ \rho}{\eta}\ [-]$$

energy requirement

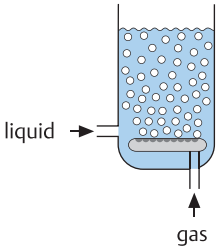
power number Ne

$$Ne = \frac{\text{propulsion}}{\text{inertia}} = \frac{P_o}{d_R^5 n^3 \rho}\ [-]$$

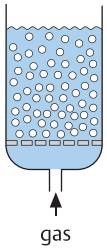
d_R = impeller diameter [m]
n = impeller revolutions [s^{-1}]
η = dynamic viscosity [Pa • s]
P_o = impeller performance [W]
ρ = density [kg/m^3]

Aeration

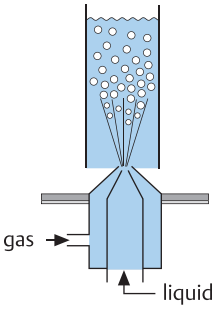
sparger



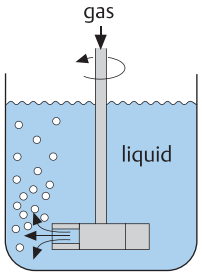
sieve plate



injector



self-aspiring stirrer



specific O_2 requirement Q_{O_2}

$$Q_{O_2} = X q_{O_2} = k_L a\ (C_{O_2}^* - C_{O_2})\ [mol \cdot L^{-1} \cdot h^{-1}]$$

$q_{O_2}^{max}$ = maximum O_2 resorption rate [$mol \cdot g^{-1} \cdot h^{-1}$]
q_{O_2} = specific O_2 resorption rate [$mol \cdot g^{-1} \cdot h^{-1}$]
$k_L a$ = volumetric transfer coefficient [h^{-1}]
X = biomass concentration [$g \cdot L^{-1}$]

O_2 resorption rate r

$$q_{O_2} = q_{O_2}^{max}\ \frac{C_{O_2}}{K_{O_2} + C_{O_2}}\ [mol \cdot g^{-1} \cdot h^{-1}]$$

C_{O_2} = dissolved O_2 concentration [$mol \cdot L^{-1}$]
$C_{O_2}^*$ = saturation concentration O_2 [$mol \cdot L^{-1}$]
K_{O_2} = Michaelis constant O_2 [$mol \cdot L^{-1}$]

O_2 transfer: $k_L a$ value

$$k_L a = k \left(\frac{P}{V_R}\right)^{\alpha} (u_G^0)^{\beta}\ [h^{-1}]$$

k, α, β = constants [–]
p = impeller performance [W]
V_R = reactor volume [m^3]
u_G^0 = gas superficial velocity [m/s]

aeration rate B

$$B = \frac{\text{volume air}}{\text{bioreactor volume}}\ \text{min (vvm)}$$

aeration number N_B

$$N_B = \frac{V_G}{n d_R^3}\ [-]$$

V_G = gas volume current [m^3/s]
n = impeller revolutions [s^{-1}]
d_R = impeller diameter [m]

Fermentation technology: scale-up

General. During the transfer of processes from a development to a production scale (scale-up), changes in several parameters must be considered. Depending on the process and the desired production volume, several types of bioreactors can be used, the stirred reactor being the most popular version. Traditionally, scale-up is done in decimal steps (30 L to 300 L to 3000 L to production scale).

Scale-up. Even on a pilot scale, bioreactors are equipped with impellers, turbines, baffles, pumps, or aeration modules to allow good mixing. For results at the pilot scale to be transferred to production scale, one must take into account that mixing time strongly increases with volume and that fast mixing in reactor volumes > 150 m^3 is not only difficult to achieve, but entails prohibitive energy costs. This fact also plays against plasmids that carry a λ promoter: the rapid rise in temperature required for induction (→62) is completely out of scope in a large production fermenter. Apart from mixing, mechanical shear stress also sets limits to fermentations using *Streptomyces* strains or fungal strains such as *Aspergillus* or *Penicillium*. Similar arguments are valid for the distribution of air bubbles during aeration as well as for heat removal.

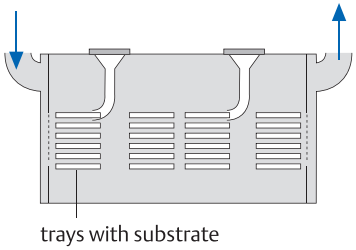
Bioreactor types. Solid state bioreactors (→86), e. g., for the manufacture of citric acid, are mostly of historical importance, although the trickling filter in aerobic wastewater treatment is still a widely-used example of this reactor type, and solid-state fermentation (SSF) is becoming popular again in the context of waste use. Surface reactors are simple to operate but their volume-space-yield of products is rather limited. For industrial products, stirred bioreactors are the equipment of choice. They are temperature- and pH-controlled, equipped with stirrers and aeration systems, and have sterile feed and sampling valves. Stirrers are usually of a multi-stage type, complemented by baffles; in some reactors, one-stage stirrers with overflow and conduit tubes are still being used. For acetic acid production (→144) and aerobic wastewater treatment (→286), self-aspiring agitators are in use. Research-type bioreactors range from 1 to ca. 300 L, and stirred bioreactors with a working volume up to 500 m^3 can be found in industrial manufacturing. For the optimization of inocula, media etc. ("upstream bioprocessing"), smaller volume parallel bioreactor systems, which provide scalable results, have become popular. For upstream experiments or for fermentation experiments at a smaller-scale, disposable bioreactors have become popular. They are supplied as, e. g., sterile plastic bags. At larger volumes, energy requirements for fast mixing and heat transfer quickly increase. Thus, loop or airlift reactors are preferred for large-volume fermenters with a volume up to 1500 m^3; in these systems, mammoth pumps or injectors serve as mixing devices. Single-cell protein manufacture (→122) and aerobic waste water treatment (→286) have been mentioned in this book as typical examples in which airlift reactors are being used.

Measurement and control are most important for the economic optimization of a bioprocess, but also for the operational safety of a plant. Routinely measured parameters include reactor weight, temperature, pH value and O_2 content of the nutrient broth, and revolutions and energy use of the stirrer. Usually, feed and exhaust air are analyzed for CO_2 (by IR spectroscopy) and O_2 (by paramagnetic resonance), or both by mass spectrometry, providing the respiratory quotient (RQ), which yields important clues about the growth and condition of a culture. Consumption of substrate and formation of product are usually determined outside the bioreactor, after sterile sampling. In view of the high value of a bioreactor filling (even if based on a market value of just 10 € kg^{-1} product and an end-product concentration of 100 g L^{-1}, the value of product after fermentation in a 100 m^3 bioreactor is 100,000 €), methods for reliable operational and sterility control are continuously being updated.

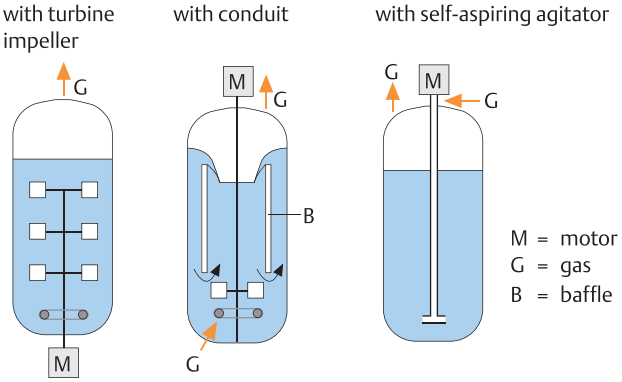
Foam breakdown. Foam forms during aeration of proteinaceous or surfactant solutions and is a frequent impediment in fermentation processes. It is usually counteracted by a mechanical foam breaker (thermal foam destruction or foam centrifuge), which is located on top of the impeller shaft. If foam formation is very heavy, chemical antifoam agents such as erucic acid or silicones have been used. Their disadvantage is, however, that they may wind up in the end product of fermentation and may be difficult to remove.

Bioreactor types

surface reactor

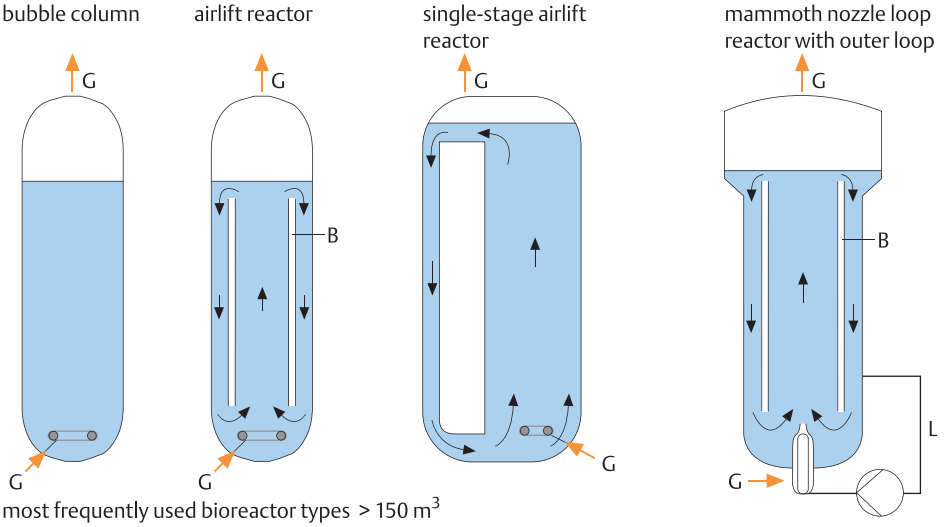


stirred reactors



most frequently used bioreactor type up to ~150 m^3 volume

bubble, airlift, and shaft reactors



most frequently used bioreactor types > 150 m^3

Mixing time – dependence on bioreactor volume

volume [L]	3	9	100	300	1,000	3,000	24,000
impeller speed [Upm]	750	2,000	230	350	200	180	30
mixing time [s]	5	3	6.6	5	25	20	66

Measurement and control in bioreactor technology

physical parameters	chemical parameters	biological parameters
temperature	**pH value**	enzyme activities
pressure	dissolved oxygen	ATP content
power input	O_2 and CO_2 in exhaust gas	NADH content
viscosity	redox potential	protein content
flow rate of air	**substrate concentration**	
nutrient feed	product concentration	
turbidity	ion concentration	
bioreactor weight		

measured parameters **regulated parameters**

Cultivation of mammalian cells

General. Mammalian cell cultures are preferentially used 1) to produce vaccines (→248, 250), and 2) to manufacture therapeutic and diagnostic proteins that cannot be obtained from recombinant microorganisms. These include proteins that contain numerous disulfide bridges, that are effective only after complex posttranslational modifications (e. g., glycosylations) (→262), or which, due to a wrong glycosylation pattern, lead to an immune response upon long-term administration. Examples are therapeutic antibodies (→246), factor VIII (→228), erythropoietin (→238), and tPA (→230). The manufacture of recombinant proteins in animal cell culture is technologically demanding and expensive. As an alternative technology, attempts are being made to produce these type of proteins in transgenic animals or plants. Human tissue cultures are also being investigated for the manufacture of tissue replacements for transplantation medicine (→308) or for the testing of drugs and cosmetics (tissue engineering).

Human cell cultures. Cell populations taken from human tissue and expanded in nutrient media have been used for a long time for the identification and propagation of human-specific viruses. Human cells can be stored in the gas phase of cryocontainers at −196°C, and uniform cell cultures can be obtained from such stocks (master cell banks) over a long period of time. The life span of these primary cells, however, is limited to ca. 50 cell divisions, and they absolutely require a solid surface for growth. As a result, the yield of cell material obtained by this method is limited. In contrast, continuous cell lines have the capacity to expand (propagate) indefinitely. Examples of continuous cell lines are HeLa cells and the Namalva cell line, obtained from human cervix cancer and lymphoma, respectively. Embryonic stem cells (ES), although not cancerous, are also continuous cells, since they can divide indefinitely. Both ES and induced pluripotent stem cells (iPS) obtained from fully differentiated cells of a patient are currently under intense investigations (→306), as they can be differentiated into a wide range of organ-specific cell types and are not rejected by the patient upon transplantation since they originate from his own body (autologous cells). They are also very interesting sources for tissue engineering (→308) –autologous cell cultures, often from different cell types, can be used in tissue replacement therapies such as for damaged cartilage, bone or "artificial skin" for burns.

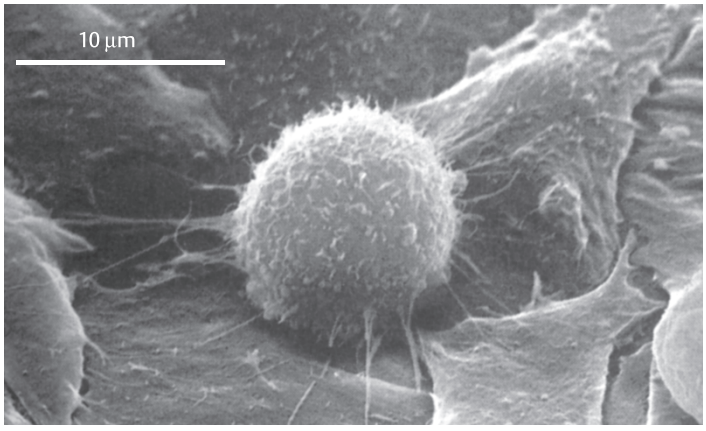
Cell lines in biotechnology. Continuous cell lines are useful for the production of therapeutic proteins. Cell lines used in industry originate from animal tumors. They divide without limit and exhibit these properties: 1) short doubling time of 20–30 h; 2) uncomplicated culture conditions; 3) ability to grow in suspension to a high cell density with adequate stability to shear stress, which allows their cultivation in large bioreactors; and 4) availability of vectors for transformation. Today, the following cells are mainly used for this purpose: 1) mouse or human hybridoma cells for the manufacture of humanized or human monoclonal antibodies, 2) Chinese hamster ovary fibroblasts (CHO cells); and 3) tumor cells from Syrian baby hamster kidneys (BHK cells). CHO or BHK cells are widely used for expressing recombinant proteins, e. g., tPA, EPO, or factor VIII. They lead to products whose post-translational modifications, in particular their glycosylation patterns, are very similar to the native human proteins.

Cloning vectors. Several types of vectors are used for preparing genetically stable recombinant animal cells. They all integrate into the genome of the target cell and may carry a sequence for external induction. Usually, shuttle vectors are used that can also transform *E. coli*, which is used as a host to optimize the vector. In laboratory experiments, typical markers applied for the selection of transformed cells are proteins conferring resistance to toxic medium components such as neomycin or Cd salts. For industrial applications, such agents cannot be used. Instead, dihydrofolate reductase (DHFR) in combination with *dhfr*-deficient host cells (e. g., CHO-K1) (→62) is the marker of choice. DHFR is competitively inhibited by the folic acid analog methotrexate, disturbing thymidine biosynthesis. Thus, transformed *dhfr*-deficient cells auxotrophic for thymidine can grow on a minimal medium in the presence of metothrexate after transfection with a DHFR vector. They even amplify the foreign *dhfr* gene, ligated to the cDNA of the desired gene product. Two vectors used for the industrial production of factor VIII (→230) and tPA (→232) in CHO cells are shown on the opposite page.

Frequently used permanent cell lines

	origin	applications
Namalva	human lymphoma	interferon, factor VIII, tPA, EPO, FMD virus vaccine, monoclonal antibodies, human vaccines etc.
BHK	baby hamster kidney	
CHO	Chinese hamster ovary	
SP 2/0 mouse hybridoma	mouse myeloma	
BS-C-1 vero cells	primary monkey kidney	

Animal cells used for protein production

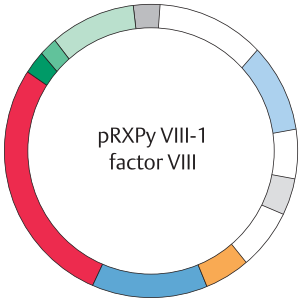


CHO cell (attached to a surface)

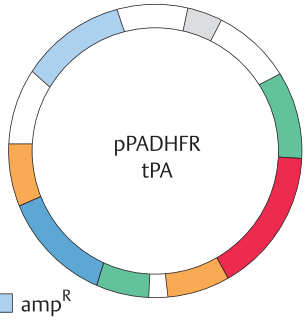
desired properties

- immortal
- simple and stable transformation
- propagate to high density in suspension culture
- high duplication rate
- undemanding culture conditions
- high specific productivity

Examples of CHO expression vectors



- colE1 ori
- tet resistance gene
- polyoma virus (ori/enhancer)
- adenovirus (late promoter)
- first exon adenovirus (late genes)
- following intron with 5- and 3-splicing recognition site
- factor VIII cDNA
- dhfr-cDNA
- SV40 (early terminator)



- amp^R
- colE1 ori
- SV40 (early promoter)
- tPA cDNA
- HbsAg (terminator)
- dhfr cDNA

Selection and amplification markers

marker gene on expression vector	component in nutrient medium	selection/ amplification	comments
neomycin phosphotransferase	neomycin	selection of neomycin-resistant cells	unsuitable for production media
metallothionein	Cd ions	selection of Cd-resistant cell lines	research use
dehydrofolate reductase (DHFR)	thymidine and methotrexate (a DHFR inhibitor)	selection of dhfr-complemented cell lines, amplification by methotrexate	frequently used for production

Mammalian cell bioreactors

General. In recent decades, numerous protocols for the cultivation of mammalian cells on a laboratory scale have been developed. In particular, nutrient media have been optimized. Because several recombinant proteins such as human or humanized antibodies (→246), factor VIII (→230), and some forms of tPA (→228) can be produced in sufficient quality and quantities only in mammalian cell culture, manufacturing processes based on mammalian cells have been scaled up to 20,000 L. In this context, typical aspects of bioprocess engineering such as mixing and aeration play a key role (→94). Purification protocols for animal cell products are extremely demanding, in view of excluding undesired protein constituents and even traces of viral DNA or retroviral RNA.

Nutrient media. Besides a good supply of oxygen and CO_2, a sufficient supply of nutrients is of utmost importance. Unlike microorganisms, animal cells grow best within a narrow temperature and pH regime (usually 37 °C, pH 7.0). To regulate cellular pH, bicarbonate is added, or experiments are conducted under a CO_2 atmosphere. Glucose is preferred as the carbon source. Amino acids, vitamins, nucleotides, proteins, fatty acids, inorganic salts, and growth-promoting substances must also be included. Most older nutrient mixtures contained fetal calf serum as a complex source of growth promoters. The serum-free media used today are less expensive and more acceptable with respect to animal welfare concerns. They contain supplements of insulin, transferrin, ethanolamine, and selenite (e. g., the ITES medium); technical scale preparations also contain plant-derived peptones and lipoproteins.

Laboratory procedures. In the laboratory, cells are preferentially grown in roller or T flasks (tissue flasks). Up to 10 L, spinner flasks are used, where cell concentrations can reach 1–2 million cells mL^{-1}.

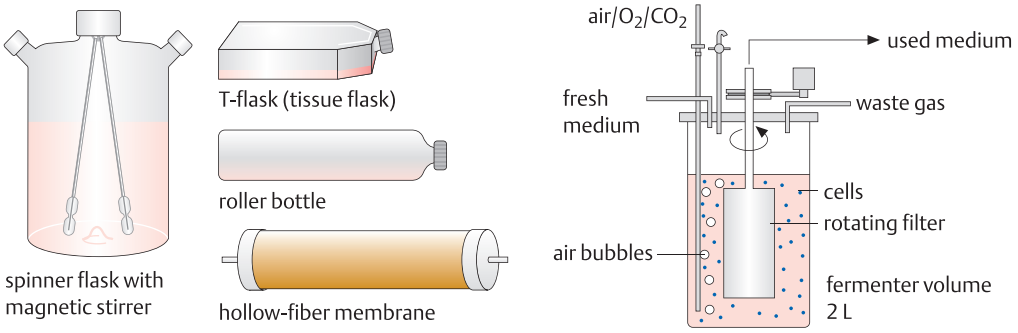
Cell reactor. For scaling up experiments to a cell reactor, one must be aware that toxic metabolites may enrich during batch or fed-batch cultivation and inhibit formation of the desired product. In perfusion cultures, spent medium is replaced with fresh nutrient solution. For industrial fermentations, however, fed-batch protocols are preferred (→92), since they reduce the risk of infection. Because cells are in suspension, oxygen supply and shear stress stability are important but opposing requirements that must be optimized. The k_La values for the oxygen supply of animal cells are on the order of 2.2 h^{-1}, or 1–2 magnitudes lower than the k_La values for microorganisms (→94). A wide range of indirect oxygen supply systems has been investigated, e. g., semipermeable silicon membranes. For example, in a 1000-L bioreactor with a rotating membrane system, an oxygen transfer efficiency of 120 g O_2 m^{-3} was achieved, applying an internal membrane pressure of up to 6 bar. Another important task is full exploitation of the expensive nutrients. In perfusion experiments, the cells are retained by a microporous membrane filter. Once the desired cell density is reached, continuous workup of the product begins. Cultures of this type have been held at equilibrium for > 900 h. With cells fixed on microbead carriers with a large internal surface, e. g. made from porous silicon, a stable production yield was obtained for over 30 d. Modern production techniques up to the 20,000 L scale need about 14 days to complete. Only fed-batch procedures are used, and cells are suspended in a serum-free medium. At smaller scales (up to 100 L), continuous perfusion modes are used where spent medium containing inhibiting metabolites such as lactic acid and ammonia is replaced. The cells are filtered from the harvest stream by a microporous spin filter and recycled. Once a cell density of 15–20 mill. cells/mL is reached, product recovery begins. Cells grown on microparticles are also used, which can be kept in equilibrium for up to 30 days.

Purification of products. In contrast to most microbial products, recovery of mammalian-cell products does not require breaking the cells, since the desired products are secreted into the culture medium. However, lengthy purification protocols must be followed to achieve highly pure products, and quality control is a key issue. Recovery steps may include affinity purification. Identity of the protein product is confirmed using peptide maps, terminal amino acid sequencing, MALDI-TOF, and other methods; the absence of oncogenic or transformable DNA must be proven to a concentration of 100 pg (10^{-10} g) per dose of the pharmaceutical end product, and no retroviral RNA or endotoxins can be present.

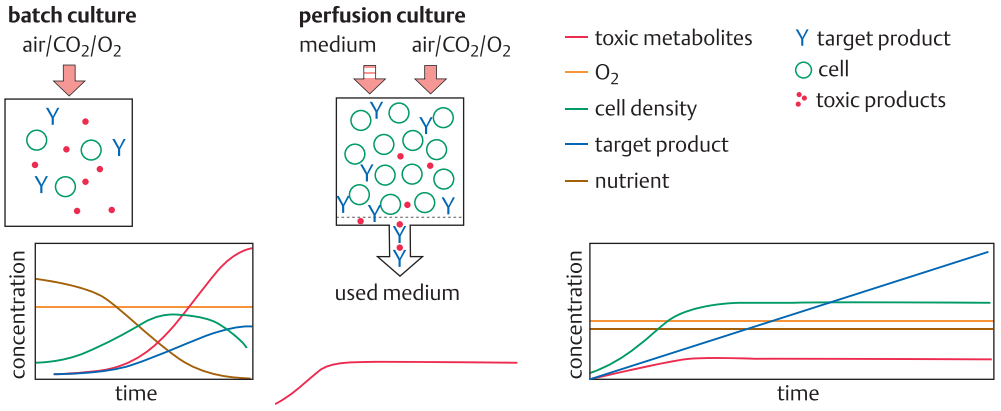
Nutrient media

serum-containing media	serum-free media
glucose, glutamine	glucose, Na pyruvate, glutamine
amino acids	amino acids
fetal calf serum	insulin, transferrin, ethanolamine, selenite (ITES medium)
salts	salts
CO_2- and O_2 supply	CO_2- and O_2 supply

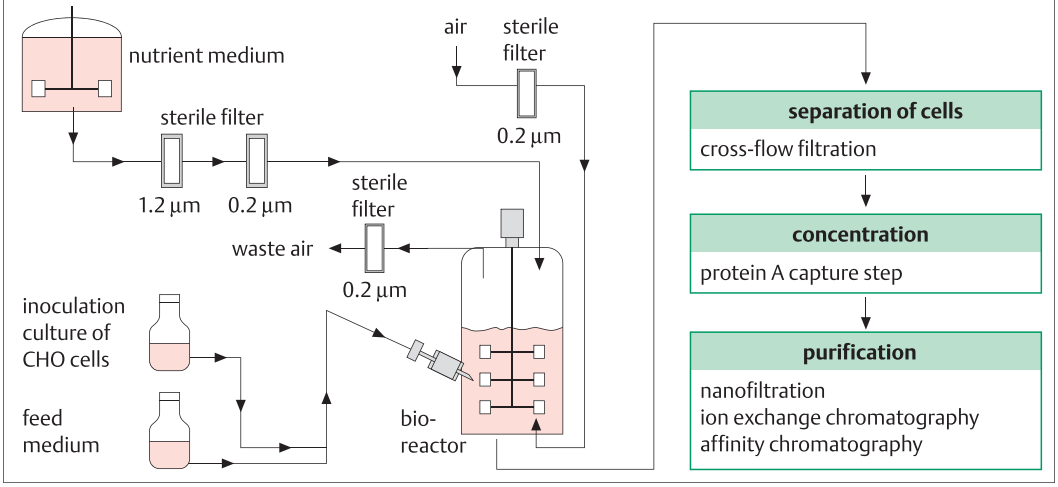
Cultivation in the laboratory and pilot plant



Methods of cultivation



Scheme of a fed-batch culture using CHO cells



Enzyme and cell reactors

General. With genetic engineering, the manufacture of enzymes has become much simpler and less expensive. More recently, whole microbial cells have been optimized for the production of a single desired product using metabolic engineering (→318) methods. Industrial reactions with enzymes or whole cells are either done with the dissolved or suspended catalyst or, both for stability and cost, with enzymes or cells immobilized on a carrier (→164). Immobilization on or within a carrier material can be achieved by adsorption, inclusion or covalent binding. Examples for industrial processes with immobilized enzymes are the cleavage of penicillin G to 6-aminopenicillanic acid (→208), and the lipase-catalyzed manufacture of chiral amines (→170). Immobilized bacterial cells are used for the synthesis of L-aspartic acid from fumaric acid (→103), and for glucose isomerization (→180), where the cells producing the appropriate intracellular enzyme (glucose isomerase) are inactivated under conditions where the desired enzymatic activity is preserved.

Chemistry of immobilization. Adsorption of enzymes or cells onto charged surfaces may often suffice: to this end, a wide range of ion-exchange or micro- and mesoporous materials are available. If covalent immobilization of an enzyme is preferred, 3 main methods have been used: 1) crosslinking of surface ε-amino groups of lysine with glutardialdehyde, leading to azomethines which can be further stabilized by hydrogenation with sodium borohydride; 2) crosslinking with diisocyanates; and 3) binding to polymeric epoxides (oxiranes). Numerous inorganic and organic carriers have been investigated as matrix materials. For the inclusion of cells within a carrier material, prepolymers are mixed with cells, and the mixture is subjected to radical or photochemical polymerization. For example, cells (or enzymes) can be incorporated into a polyacrylamide gel when they are mixed with acrylamide, and then N,N'-methylene bisacrylamide and potassium persulfate are added. Urethane prepolymers and other materials have also been used for photochemical polymerization in the presence of enzymes or cells. For microencapsulation of enzymes or cells, polymerization of suitable prepolymers may be carried out at the boundary phase of water, which contains the enzyme or cells, and an organic solvent that does not mix with water. Water-soluble polymers such as polyethylene glycol may be used for immobilization of detergent enzymes. Another popular method uses ionotropic gels such as alginate (→18), which form a gel in the presence of Ca^{2+}. Weaving of enzymes and cells into fibers, e. g., collagen, has also been described.

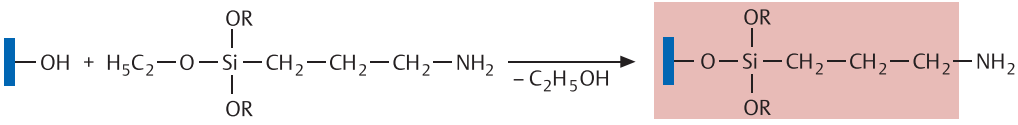
Properties of immobilized biocatalysts. Upon immobilization, biocatalysts may change their properties, because catalytic efficiency is not influenced only by the catalyst itself, but also by mass transfer, which depends on the properties of the immobilization matrix used. Thus, empirical rules have been elaborated that are useful for optimizing the immobilized catalyst as well as the packing of bioreactors and their volume-time yields. Besides the intrinsic properties of the catalyst such as V_{max} and K_M, these rules take into account both particle size and mass-transfer parameters. A key property of an immobilized biocatalyst is its operational stability. In favorable cases, it may remain active for several months (aspartase, glucose isomerase, penicillin acylase).

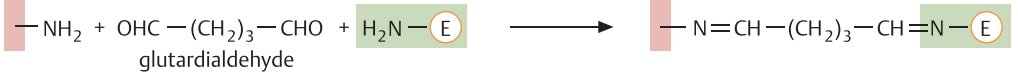
Reactor type and process technology. Enzyme and cell reactors have been engineered as discontinuous or continuous stirred reactors (→94); as fixed-bed, fluidized-bed, or hollow-fiber column reactors, and as membrane reactors. In addition to biocatalyst stability and optimization of mass transfer, upstream and downstream processing are most important for good economy. These processes include preparation of substrates, suppression of side reactions, and development of a recovery protocol (→104) that is adequate for the intended use of the product. The overall scheme must be robust, simple to operate, and optimized for minimal investment and operational costs. Measurement and control are also key issues. For large-volume products such as isomerose (HFCS) or 6-aminopenicillanic acid, continuously operating plants are preferred. For this, several bioreactors are used in parallel but with consecutive catalyst fillings; thus, each module with exhausted catalyst activity can be exchanged without decreasing the overall productivity of the plant. For products with lower tonnage such as L-aspartic acid, a single-cell reactor of immobilized *E. coli* may suffice for a production run.

Materials used for immobilization

inorganic carriers		natural and synthetic polymers		
Al_2O_3	$Ca_3(PO_4)_2$ gels	activated charcoal	oxiranes	collagen
bentonite	porous ceramic materials	polyacrylamide	agarose	polyamide
porous glass		carboxymethyl cellulose	alginate	cellulose
		polyurethanes	dextrans	

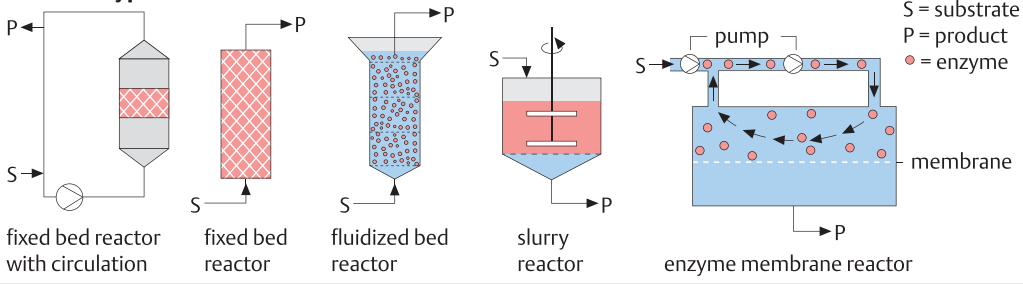
Immobilization onto glass surfaces





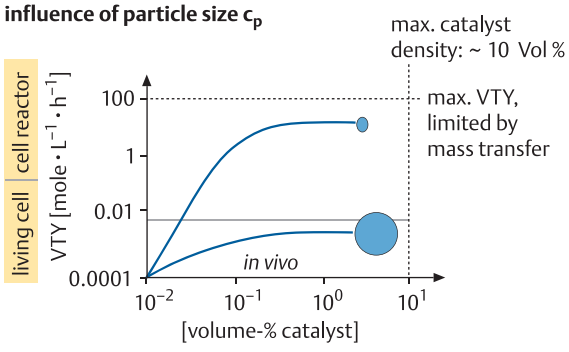
enzyme with ε-amino groups of lysine glass surface

Reactor types



Volume-time yields (VTY)

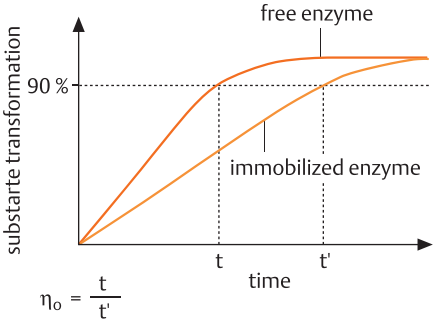
influence of particle size c_p



$$VTY = \frac{S_o - S_t}{t} \quad [mole \cdot L^{-1} \cdot h^{-1}]$$

S_o = starting concentration of substrate [mol · L^{-1}]
S_t = concentration of substrate at time t [mol · L^{-1}]

influence of enzyme immobilization



$$\eta_o = \frac{t}{t'}$$

η_o = degree of utilization of catalyst
t = time for defined substrate turnover for free (t_{12}) or immobilized (t'_{12}) biocatalyst

Influence of mass transfer

inner mass transfer: **Thiele modulus** φ

$$\varphi = \frac{\text{reacting substrate}}{\text{diffused substrate}}$$

$$\varphi = \frac{d_p^2}{4} \cdot \frac{\delta_{max}}{K_m \cdot D^e}$$

external mass transfer: **Sherwood number** Sh

$$Sh = \frac{\text{mass transfer}}{\text{diffusion}} \qquad Sh = \frac{k_f \cdot d_p}{D_0} \quad [-]$$

D^e effective diffusion coefficient [$m^2 \cdot s^{-1}$]
k_f mass transfer coefficient [$m \cdot s^{-1}$]
d_p particle diameter [m]
D_0 molecular diffusion coefficient [$m^2 \cdot s^{-1}$]
δ_{max} maximum catalyst density
K_m Michaelis constant

Recovery of bioproducts

General. Bioproducts obtained during a fermentation process are either cellular (cell mass, intracellular proteins, inclusion bodies) or extracellular (amino acids, antibiotics, enzymes). In classical fermentations, the product concentration is often low (< 10 %, often < 1 %). Using genetically engineered microorganisms, higher product yields are usually obtained (e. g., up to 50 % protein per wet cell mass). For isolating purified bioproducts, down-stream processing (DSP) is an important and often cost-determining step in manufacturing. The sequence of concentration and purification steps should be appropriate for the intended use of the product. Thus, pharmaceutical and diagnostic preparations require complex purification protocols, but technical enzymes are enriched using fewer and simpler processing steps. Product loss during purification may exceed 50 %, indicating the importance of good protocols for economic production. Safe and inexpensive disposal of waste fractions is another task of economic and ecological importance.

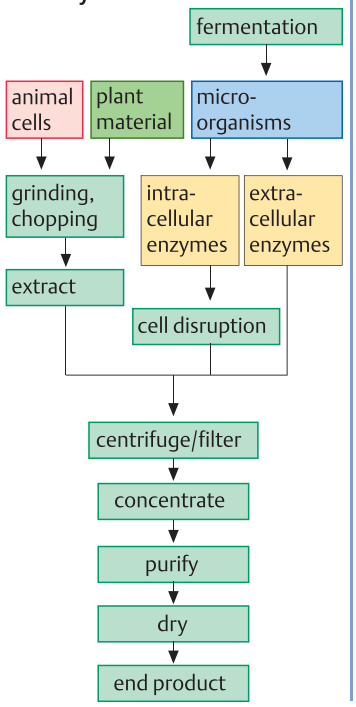
Cell mass. The manufacture of baker's yeast is a good example of the preparation of microbial biomass (→120). After fermentation is complete, cells are separated using centrifuges or centrifugal separators, followed by washing and filtration through a pressure leaf or vacuum rotating filter; the moist product is then packaged. Dried baker's yeast is produced from moist yeast with a cyclone and can be stored much longer. Diafiltration, cross flow filtration or two-phase extraction systems are additional methods sometimes used in cell separation.

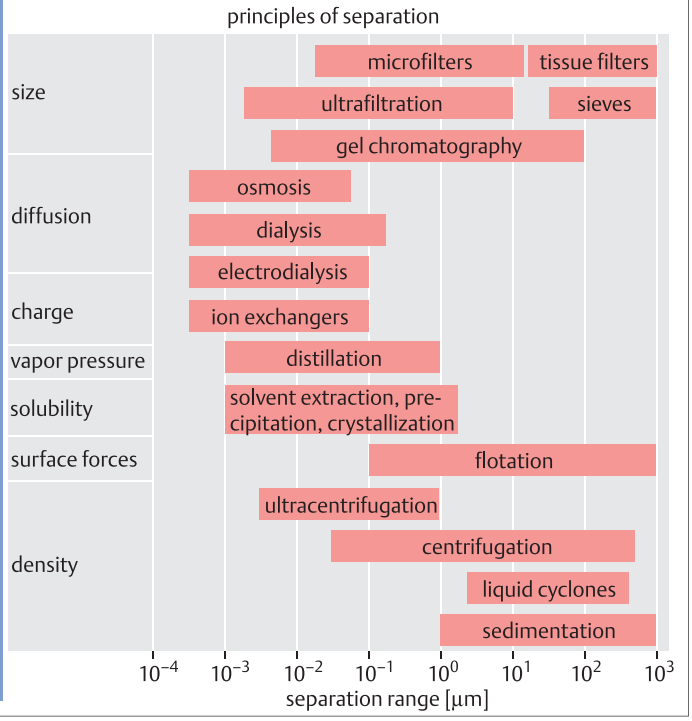
Intracellular products. The target products are usually intracellular proteins, and the separated cells must be disrupted without inactivating proteins. Mechanical methods such as high-pressure homogenizers or ball mills are often used for this step. For the production of yeast autolysate, intracellular proteases of yeast are activated and lead to cell destruction. On a laboratory scale, ultrasonication in cooled baths is often used. Lysozyme or other lytic enzymes can also be used in combination with mild surfactants. During fermentations using recombinant *E. coli* cells, target proteins often form inclusion bodies, in which disulfide bridges have been wrongly formed. Depending on the leader sequences used, inclusion bodies may form in the periplasmic space or in the cytosol. They are isolated after mild cell breakage by differential centrifugation. Reduced forms of inclusion bodies (from the cytoplasm) are first oxidized by oxidative sulfitolysis; the resulting S-sulfonates are then reduced with thiol reagents (as are oxidized forms of inclusion bodies from the periplasmic space), and the proteins are denatured with urea or similar reagents that break hydrogen bonds (→222). In a following dialysis step under oxidizing conditions, urea is removed, and part of the protein folds into the correct conformation. Overall yields of functional protein rarely exceed 20 % with this process.

Extracellular products. Low-molecular-weight bioproducts such as citric acid (→146), lactic acid and amino acids (→124) are often precipitated from the broth after removal of the cell mass. They are purified further by dissolution and precipitation steps, often leading to crystalline products. As such procedures lead to large amounts of by-products which must be disposed of, the development of novel DSP procedures based on the use of, e. g., liquid ion exchangers, are of high economic significance. Antibiotics are usually isolated by multi-stage solvent extractions (→208) using n-pentanoyl- or n-butyl acetate. The isolation of extracellular proteins is usually initiated with an ultrafiltration step, followed by salting out with ammonium or sodium sulfate. Alternatively, low concentrations (2–10 %) of cooled organic solvents such as 2-propanol can be used. Technical enzymes are often used in crude form, either as liquid concentrates or in spray-dried form, whereas diagnostic or therapeutic enzymes must usually be purified to homogeneity.

Integrated procedures. Numerous attempts to simplify workup protocols have been made. In expanded-bed chromatography, cell separation and product concentration occur simultaneously, by using a density gradient within a fluidized bed of ion-exchange or affinity adsorbents. Another procedure is based on two-phase systems and consists of an aqueous salt phase and a second phase composed of water-soluble polymers. Since these phases do not mix, they allow cell fragments and proteins to be enriched in different phases and have been fine-tuned for integrated product recovery.

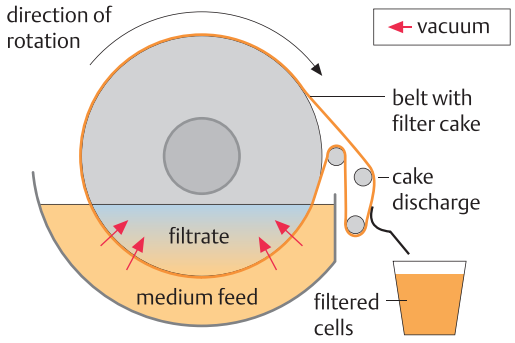
Survey



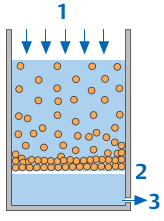


Filtration

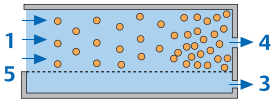
vacuum rotary filter



static filtration

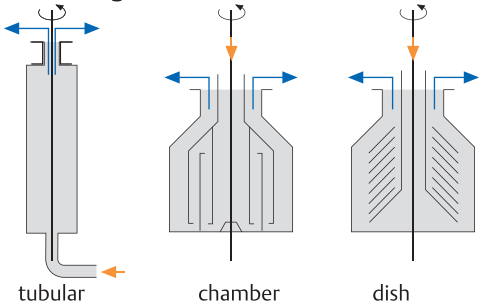


cross-flow filtration



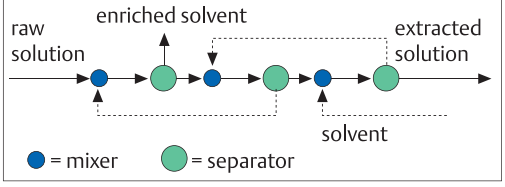
1 culture broth
2 filtration aid
3 filtrate
solid particle
4 concentrate
5 membrane

Centrifugation



Solvent extraction

multistage counter-current extraction



Membrane-based procedures

	reverse osmosis	ultrafiltration
principle	transport by diffusion	separation by molecular size
retained particles	M_R < 500 –1 000	M_R > 1 000
osmotic pressure	0.8 – 10 mPa	very small
working pressure	1 – 15 mPa	< 1 mPa

Recovery of proteins: chromatography

General. Chromatography is a very important step in most purification protocols. The optimal layout of a chromatographic step can be calculated with the van Deemter equation. Column filling materials are classified according to the separation principle: 1) in gel chromatography, substances are separated by their molar masses and shapes; 2) in absorption chromatography, hydrophilic or hydrophobic interactions predominate; 3) in ion-exchange chromatography, charged amino acid side chains are of key influence; 4) in chromatofocusing, the isoelectric point of a protein governs separations; and 5) in affinity chromatography, separation occurs by specific interactions with ligands. For each of these methods, a wide variety of commercial adsorbents is available. A valuable piece of equipment for selecting the most suitable adsorbent and elution protocol within a few hours is the ÄKTA™ system. In the laboratory, many chromatographic separation steps are done at medium or high pressure (e.g., FPLC = fast protein liquid chromatography). Procedures scaled to the production level, however, do not include high-pressure chromatographic steps.

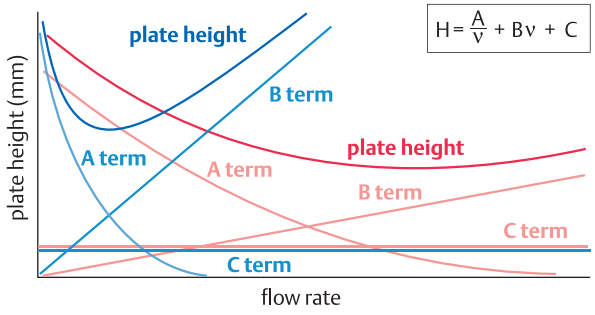
Gel chromatography. Important commercially available carrier materials are gels formed from dextrans (→158) or agarose, whose pore size can be modulated by cross-linking. After partial alkylation of hydroxyl groups by alkylating reagents, they can also be used with organic solvents; after attachment of charged groups, they combine the properties of a "molecular sieve" and an ion exchanger.

Adsorption chromatography. A hydrophilic material in wide use is hydroxyapatite. Hydrophobic chromatography is usually carried out with Sepharose gels that are derivatized with butyl, octyl, or phenyl groups. They enable purification of hydrophobic proteins through their interaction with the hydrophobic adsorbent material.

Ion-exchange chromatography. Ion exchangers are frequently used in protein purification, since they are efficient and can be scaled up easily. Sulfonated or carboxylated polymers are often used as cation exchangers. Anion exchangers are often based on polymers containing quaternary amino groups. Polysaccharides or synthetic polymers are used as a polymeric support. The separation principle rests on the difference in net charge of different proteins. Elution is done by raising the salt concentration or changing the pH.

Affinity chromatography. This elegant method relies on the interaction of proteins with a specific ligand that is coupled to the carrier material. For purification of dehydrogenases, for example, dextran-coupled pigments that fit into the NADH binding pocket of these proteins have been used. Immunochromatography (→82), a very expensive technique, is occasionally used in industrial purification of pharmaceutical proteins such as factor VIII. The carrier is loaded with monoclonal antibodies specific for this protein. Elution from the matrix is achieved with low-molecular-weight competitive ligands, or the column is eluted by raising the salt concentration or lowering the pH. Often, on the small scale, recombinant proteins are purified by using affinity ligands that have been purposely introduced into the protein by genetic engineering. Fusion proteins with an easily purified helper protein such as streptavidin (purified over a carrier carrying biotin ligands) is one example. Another useful technique is the addition of a polyhistidine tail (his_n, $n = 4–6$) to the N or C terminus of a protein. Proteins modified with such "his tags" can be purified with high selectivity on a metal-containing carrier matrix (IMAC = immobilized metal affinity chromatography), and preferentially Ni^{2+}, but also Cu^{2+}, Zn^{2+}, Co^{2+}, Fe^{2+}, or Fe^{3+} are used as matrix-bound metal chelators. If cleavage sites highly specific for a protease such as factor X are introduced adjacent to the his tag, the tag can be selectively removed after purification. Other peptide tags have been genetically introduced into proteins, e.g., a sequence binding to the biotin site of streptavidin, allowing purification over a column modified with streptavidin. The fused protein or peptide tag is usually linked in a manner that allows for its removal by a specific protease after purification. A completely different approach rests in the use of "molecular imprint" materials. By immobilizing multiple functional ligands on a hydrophilic cross-linked polymer fixed at the surface of the support material, a target compound can purified with very high selectivity.

van Deemter equation



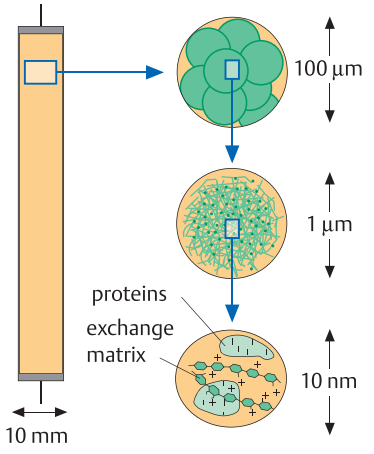
H = plate height
v = flow rate
A = diffusion term

B = term for equilibrium
C = term for column quality

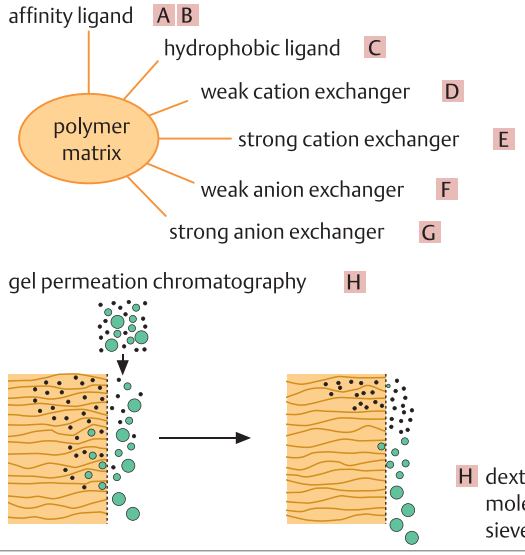
high molecular-weight compounds (proteins)
low molecular-weight compounds

Polymer matrix

example: cationic polysaccharide



Chromatography types



Some ligands

hydrophobic interaction

C butyl, octyl, phenyl

cation exchanger

D carboxymethyl
E sulfopropyl

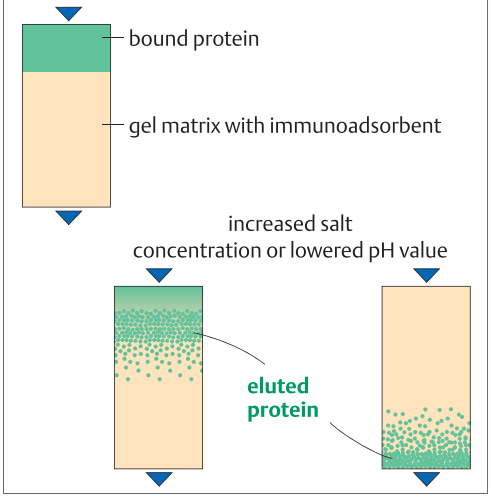
anion exchanger

F diethylaminoethyl
G trimethylaminomethyl

affinity chromatography

A group-specific ligands
B immunoreagents

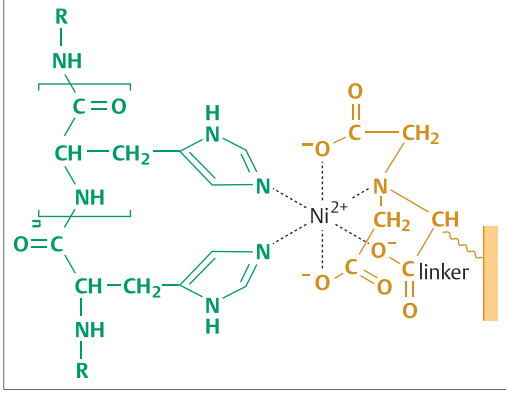
Immunoaffinity chromatography



his-tag chromatography

genetically modified protein with terminal poly-histidine tag (n = 4 – 6)

nitrilotriacetic acid (NTA) as ligand



Economic aspects of industrial processes

General. For industrial applications, biotechnological processes are optimized with economic considerations in view. The key goals are 1) to develop a simple and robust process; 2) to keep the personnel and investment costs low; 3) to obtain high yields in a short time, using inexpensive, readily available raw materials of constant quality while keeping energy costs low; and 4) to keep environmental costs (disposal of waste, wastewater treatment) low.

Simple and robust processes. During multistage processes, which are common in biotechnology, the elimination of even one process step may offer significant advantages. Thus, the so-called Westphalia decanter became very popular for production of antibiotics (→208) because it combines the two steps of cell separation and solvent extraction. For the isolation of acid bioproducts, ion exchangers are now preferred over salt precipitation as in the latter processes dangerous chemicals such as concentrated sulfuric acid were used, and large amounts of salts (gypsum) accumulated as by-products and had to be disposed of. Robust procedures are necessary because biotechnological processes take a long time and are often carried out by unskilled workers in shift operations. Infections during a bioreactor run pose a severe and permanent problem and may lead to losses of hundreds of thousands of euros or dollars.

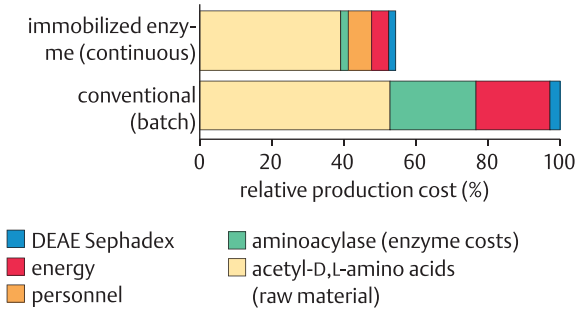
Low personnel and capital costs. Between 10 % and 40 % of biotechnology production costs are for labor. A typical fermenter run plus recovery takes about one week, and the process must be controlled by workers' shifts around the clock. The capital investment for a bioreactor production facility is on the order of 10^7 to 10^8, depending on the product. Depreciation and insurance may amount to 10 % or more of the manufacturing costs.

Energy balance and raw materials. Energy costs are mainly determined by the costs of sterilization, cooling, and stirring. The large bioreactors used in industrial processes are sterilized (→88) in a continuous mode or by steam injection (140 °C for 4 min). During cell growth, about half the energy content of the carbon source is dissipated as heat. In production fermenters, cooling jackets and helical cooling coils are used to remove this heat. With heat exchangers, ca. 90 % of the energy used for sterilization and obtained from cooling is recovered (→94). In classical fermentation processes such as citric acid production, raw materials, in particular the carbon source, may account for between 30 and 60 % of the total manufacturing cost (→88). Complex, inexpensive feedstocks are mostly used, whose composition may vary from lot to lot. Thus, standardization and quality control of raw materials is an important requirement in manufacturing.

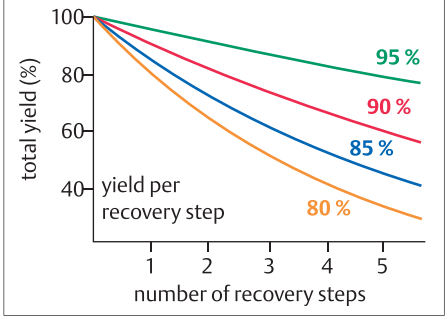
Cell mass disposal and wastewater treatment. After all fermentations, cell mass and used nutrient media accumulate in large volumes, and their BOD_5 (→286) is too high for direct disposal. For example, the production of citric acid (→146) in a 300 m^3 bioreactor leads to 15 t of wet cell mass of *Aspergillus niger* mycelium as a filter cake. For environmental reasons, these residues are incinerated – an expensive procedure for wastes with high water content. At the end of a fermentation process, the desired product is dissolved in a highly diluted aqueous solution. Even in well optimized protocols, yields rarely exceed 10 %. During product purification, nutrient-rich waste streams are produced, which may contain salts or organic solvents; their treatment in sewage plants leads to wastewater treatment costs that must be included in the price calculations.

Sensitivity analysis. Each bioprocess can be divided into single unit operations that can be subjected to economic scrutiny. This procedure helps to analyze which process steps are cost-intensive and where optimization measures will result in the largest savings. Spreadsheet programs are mostly used for these analyses, since they facilitate viewing the influence of individual factors in a networked manner. Thus, material costs for the production of recombinant human insulin immediately demonstrate that the price of the chemicals used in recovery exceeds the price of the fermentation raw materials by a factor of nearly 10, with guanidine HCl (for unfolding of inclusion bodies) and carboxypeptidase B (for cleavage of a fusion protein) being the two major cost factors. As a result, optimization of process economy will focus on protocols where these reagents are not required.

Process economics L-amino acids



Recovery losses



Aspects of recovery during industrial production of citric acid

feed solution	process step	isolated or waste product
	fermenter solution, 284,000 L	
wash water	sedimentation tank	wet biomass, 3,400 kg dry weight
	filter	
~ 22,000 kg lime [$Ca(OH)_2$]	precipitation tank, separator	calcium citrate
	heat to 80 – 90 °C, then 95 °C	
	rotary filter; in filter caken 40 000 L calcium citrate suspension	waste (filtrate)
~ 35,000 kg 95% sulfuric acid, mother liquors	acid tank	
wash water	filter	~ 6,000 kg calcium sulfate
	evaporator: 67% citric acid in residues	
	ion exchanger, activated charcoal	
	crystallizer	
wash water	centrifuge	mother liquors
hot air	dryer	
	packaging	40,456 kg citric acid, anhydrous

Material costs for production of a recombinant protein

(h-insulin in *E. coli*, old Hoechst process, 35 m^3-fermenter, batch-fed, 24 h, 25 g/L dry cell mass productivity 1 t/year)

type of cost	quantity (kg/year)	price (US$/kg)	value (US$/kg/year)	specific cost (%)
glucose	432,640	0.69	298,520	68.3
corn steep liquor	652,800	0.12	78,336	17.9
salts	~ 12,000		~ 40,000	8.4
antifoam agent	2,448	4.86	11,897	2.7
tetracycline	163.2	55	8,965	2.7
fermentation raw materials			437,000	100 rel%
guanidine HCl (for unfolding of inclusion bodies)	1 007	2.15	2,165,100	56.4
carboxypeptidase B (for cleavage of fusion protein)	0.8	1 023.00/g	818,400	21.3
formic acid	262,280	1.25	327,850	8.5
bromcyan	22,848	11.00	251,330	6.5
all other chemicals			28,000	7.3
recovery materials			3,839,000	100 rel%

Alcoholic beverages

General. Alcoholic beverages have been developed in most if not all human cultures. In Western cultures, these are primarily wine, beer (→112), fermented fruit juices, champagnes, and distilled spirits. The indigenous alcoholic beverage of Asia is rice wine (sake). Regional specialties in other parts of the world are kumys (fermented mare's milk, Mongolian nomads), kwass (from fermented cereals, Russia), pombe (from millet and sorghum, Near Orient) and pulque (from agave juice, Latin America).

Wine is produced from the fermentation of grape must or juice by yeasts. The genome of the common grapevine (*Vitis vinifera*) was decoded in 2007, and detailed genetic information now supports breeding programs. Where the vines are grown, their variety, and the production technology all play key roles in the quality of a wine. The production technology includes harvesting, pressing, must treatment, must fermentation, and cellar treatment. Harvesting depends on the weather and is critical to wine quality. During pressing, the grapes are usually removed from the stems and are squashed into must without destroying the seeds. For white wines, the must is immediately filtered and provides juice, whereas red wine must is traditionally fermented at 20 °C for 6–8 d to dissolve the red anthocyanines occurring in the grape skin in the ethanol developing from fermentation. In modern procedures, the must from red grapes is heated to 40–50 °C; after the addition of pectinases, the anthocyanines from the grape skin are dissolved within 2–4 h, and pressing can ensue. Suitable modifications of the must treatment enables the vintner to obtain (within limits) type-specific musts, even in years of poor grape quality. Depending on national legislation, sugar or acid may be added, acids may be neutralized by adding $CaCO_3$, or fermentation may be stopped by adding SO_2 or potassium pyrosulfite. Such procedures help to harmonize taste, suppress browning of the must (by inhibiting phenol oxidases), protect oxygen-sensitive pigments and aroma components, and suppress the growth of aerobic microorganisms such as acetic acid bacteria (→144), wild yeasts, and molds. Must fermentation now ensues, traditionally in wooden vats, today usually in tanks made from stainless steel or polyester materials. Inoculation may be spontaneous or through the addition of seed cultures of *Saccharomyces cerevisiae* var. *ellpsoideus* (→14). Depending on the type of must, fermentation may take from a few days to several months and can be regulated by the temperature. The residual sugar and ethanol content (7–15 vol%) of the wine can be set either by artificially stopping the fermentation or by adding must ("sweetness reserve") that has been preserved under CO_2 at a pressure of 8 bar. Dry wines contain < 9 gL^{-1} residual sugar. During the ensuing cellar treatment of the wine, chemical, biochemical, biological, and physical processes that are difficult to manage contribute to the maturation and harmonization of the wine. During cellaring, a pH > 3.2 may support the growth of lactic acid bacteria (→116); they convert maleic acid, via malo-lactic fermentation, into much weaker lactic acid and CO_2.

Champagnes are produced from quality wines by adding 1–3 % saccharose and yeast cultures. Fermentation takes place in tanks or, for the more expensive varieties, in the bottle. Champagnes must develop a pressure of at least 3 bar at 20 °C.

Spirits are produced by distillation of sugar extracts from cereals, vegetables, or fruits. Their alcohol content is between 30 and 60 %. Some of the raw materials for spirits are: wine (brandy, cognac, armagnac), sugar cane juice or molasses (arrak, rum), cereals (Korn, whisky), potatoes (vodka), fruit juices (fruit brandy), or agave juice (tequila).

Rice wine (sake). In contrast to wine or beer fermentation, sake is produced by an aerobic solid-stage fermentation (→86). In a first step, soaked and steamed polished rice (removal of the rice bran improves the sake taste) is inoculated with *Aspergillus oryzae* and incubated at ca. 30 °C. At high humidity, this leads within ca. 2 d to the formation of *koji*, a fermented material in which rice starch has been largely depolymerized into sugars. More boiled rice, water, and a yeast starter culture (*moto*) is added to the *koji* to form a mash (*moromi*), which is fermented for 20 d at 25 °C. After filtration and pasteurization *sake* is obtained. The alcohol content of bottled sake is about 15 vol%. As in other brewing processes, the quality of the starting materials (here: rice and water) are essential for the final product.

Alcoholic beverages

beer	starch from barley is degraded to sugars by barley amylases. The sugar solution is fermented by yeast to ethanol, in the presence of hop extracts
wine	grape juice is fermented by yeast
champagne	sugar and yeast is added to wine, followed by a second fermentation
cider	apple juice is fermented by yeast
sake	rice starch is depolymerized by amylases from *Aspergillus oryzae*, and sugars are fermented by yeast
whisky	extracts of barley, yeast, rye, or corn are fermented by yeast and distilled
vodka	extracts of potatoes or wheat are fermented by yeast and distilled

Production figures

beer

total (2011)	192.7 million kL
China	49.0 million kL
USA	22.5 million kL
Brazil	13.2 million kL
Russia	9.8 million kL
Germany	9.5 million kL
Mexico	8.1 million kL
Japan	5.6 million kL

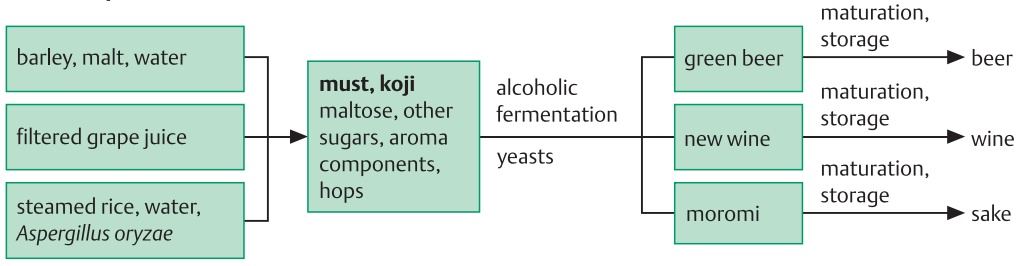
wine

total (2010)	26.4 million kL
France	4.6 million kL
Italy	4.6 million kL
Spain	3.6 million kL
USA	2.7 million kL
Argentina	1.6 million kL
Australia	1.1 million kL
Germany	0.8 million kL

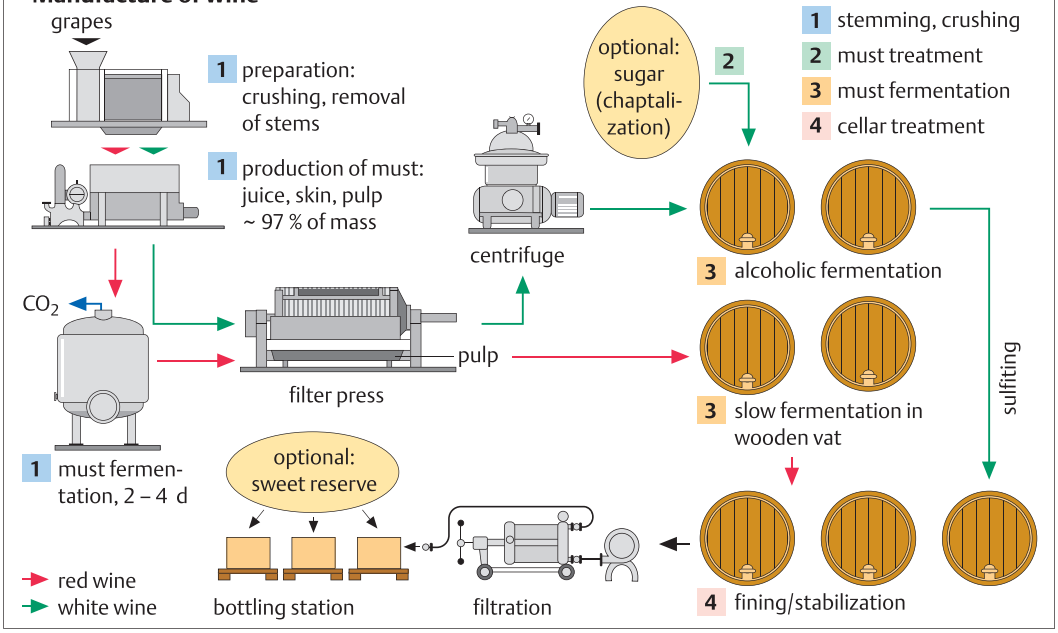
sake

Japan (2008)	0.49 million kL

General production scheme for beer, wine, sake



Manufacture of wine



Beer

General. Beer is one of the oldest and probably most widely distributed alcoholic beverages of mankind. It is produced by fermentation of malt with yeast *(Saccharomyces cerevisiae)* (→14) in the presence of extracts from hops, providing a bitter taste. Global production was about 190 million kL or ca. 190 million t. The largest producers are China, the USA, Brazil, Russia and Germany. The German "purity law" of 1516 requires that only barley (sometimes wheat), yeast, hops, and water may be used for production. In most other countries, a wider variety of raw materials may be used. Modern processing technology allows for the production of low-calorie and light (alcohol-reduced) beers. Light beers constitute over half of all beers sold in the USA.

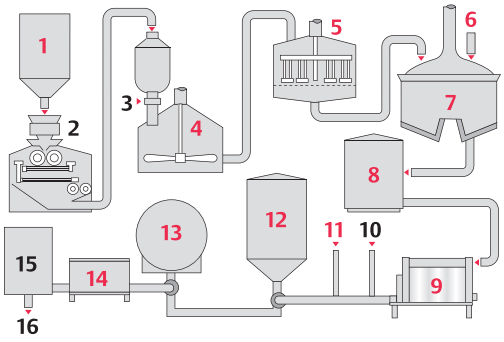
Production. Most beers are brewed with barley malt as the source of fermentable material, but other malted or unmalted starches such as wheat, rice, oats and rye are also used. By germination and kiln drying, barley is transformed into a storage-stable malt that is rich in carbohydrates and enzymes. When suspended in water, it forms a mash, to which an aqueous extract of hops is added, containing the typical bitter-taste components and balancing the sweetish taste of malt. By the action of malt enzymes, notably amylases, the resulting wort contains mono- and disaccharides, which can be fermented by yeast. The mash is inoculated with yeast cultures and fermented for some days. Depending on the fermentation process, the alcohol content of beer is between 2 and > 18 % original wort (dry weight of the extract in g/100g before fermentation). The final maturation is done in a beer cellar at temperatures around 0 °C and takes from a few days to several weeks. During this process, turbid components settle and diacetyl, an undesired off-flavor component from the fermentation process, is transformed into acetoin and butane-2,3-diol. In addition, other enzymatic transformations take place which have a strong effect on the aroma of the beer. After filtration, export-type beers are pasteurized.

Types of beer. The many varieties of beer differ according to the time and temperatures used for the main and secondary fermentations, the raw materials and their processing, and the yeast cultures used (e. g., "surface fermenting yeasts" cannot hydrolyze raffinose). Beers are mostly distinguished by the fermentation process used (bottom-fermented beers such as lager and pilsner; and top-fermented beers such as porter, ale and stout), as well as by their wort content: high-gravity beers (> 16 %), medium-gravity beers (11–14 %), draft beers (7–8 %), and low-gravity beers (2–5.5 %). In the USA, the most common type of beer is the "pale lager." For dark beers, malts with stronger colors arising from a Maillard reaction are used (heating of sugar with amino acids) (→192).

Light beers and alcohol-free beers. For the production of light (< 1.5 % ethanol) or alcohol-free beers (< 0.5 % ethanol), either fermentation is stopped early, the fermentation temperature is increased, or the alcohol is removed by appropriate processing technology after fermentation (vacuum distillation, membrane-based processes).

Biotechnological innovations. In most countries except Germany, saccharification of starch, degradation of proteins, or filtration steps are enhanced by the addition of microbial enzymes. Addition of the plant protease papain is sometimes used to prevent precipitation of proteins upon storage in the cold ("chill proofing") (→172). Recent developments target improvements such as: 1) breeding and use of transgenic barley, in particular with enhanced activity of thermostable β-glucanases and other enzymes, 2) use of lactic acid bacteria as starter cultures (→114) in malting, to minimize microbial contamination of malt by *Fusarium* species and other microbes, 3) reduction of the lengthy maturation period required to degrade diacetyl, through the addition of bacterial α-acetolactate decarboxylase, an enzyme that transforms α-acetolactate directly into acetoin, 4) the use of recombinant strains of *S. cerevisiae* which express genes such as glucanases or amylases (→172) (for better use of raw materials) or α-acetolactate decarboxylase (for faster aroma formation). Using such strains, "low calorie beer" with a very low content of carbohydrates can be brewed. Finally, the processing technology is systematically being improved. Thus, the optimization of beer taste through a week-long storage in the beer cellar can be replaced by a short period of heating to 90 °C followed by 2 hrs of maturation in a bioreactor using immobilized yeasts.

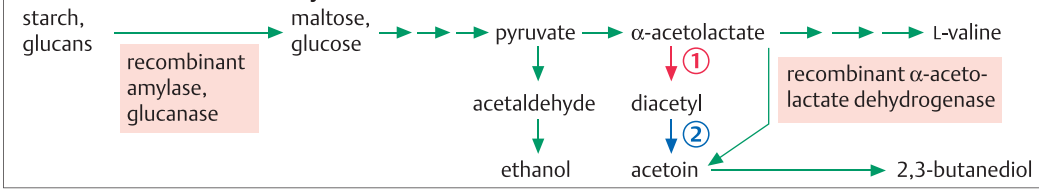
Production scheme



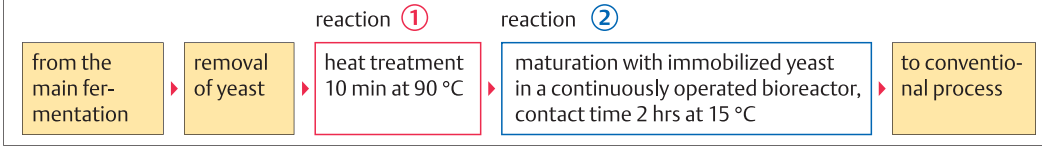
1 malt
2 mill
3 water
4 mash tank
5 lauter tun
6 hops
7 wort cooking
8 whirlpool
9 cooled filter press
10 air
11 yeast
12 fermenter
13 storage tank
14 beer filter
15 pressure tank
16 bottling station

	term	process step	details
1	steeping and malting	barley is germinated	8 d at 10 – 18 °C and pH 7.0. Amylases and proteases are activated
	kilning	barley endosperm is dried	stepwise reduction of water to < 3 % and increase of temperature to ~ 125 °C
	grain removal	barley germ is peeled	animal feed
4	mashing	dried malt is reactivated	in the mash tun, malt amylases depolymerize barley starch to dextrins and maltose, and aroma precursors are formed
5	lautering	turbid polymer material is removed	
6 7	hop and wort	hop extract is added and cooked with wort	after 1 ½ – 2 ½ hrs in the cooker, the malt enzymes are inactivated, bitter-taste compounds are dissolved, protein is precipitated and the nutrient solution is sterilized
8	filter	sediments are removed	
9	sparge	wash sediments	
11	pitching	inoculation	with pure culture of *Saccharomyces cerevisiae*
12	main fermentation	fermentation	bottom fermentation: 8 – 10 d at 5 – 10 °C ; top fermentation by yeasts and lactic acid bacteria: 2 – 3 d at 10 – 25 °C
		yeast is removed	a) decantation or b) removal of surface layer
13		beer matures	several days or weeks in the beer cellar at 0 °C. Polymers sediment, diacetyl is reduced to acetoin
14		filtration, pasteurization	lager beer

Potential of recombinant yeasts



Beer maturation using immobilized yeast



Fermented food

General. In all human cultures, many foods are modified by microbial fermentation. At first, traditional skills evolved with the aim of preserving foods. Thus, the food value of vegetables can be prolonged by lowering the pH through formation of organic acids (sauerkraut, pickles), digestibility can be enhanced through enzymatic hydrolysis prior to storage (sourdough, sausages, tempeh), taste can be improved (fermented milk products) (→116), or flavorings can be prepared (soy sauce (→86), miso). In industrialized countries, about ⅓ of the food is modified by fermentation, usually with defined microbial starter cultures.

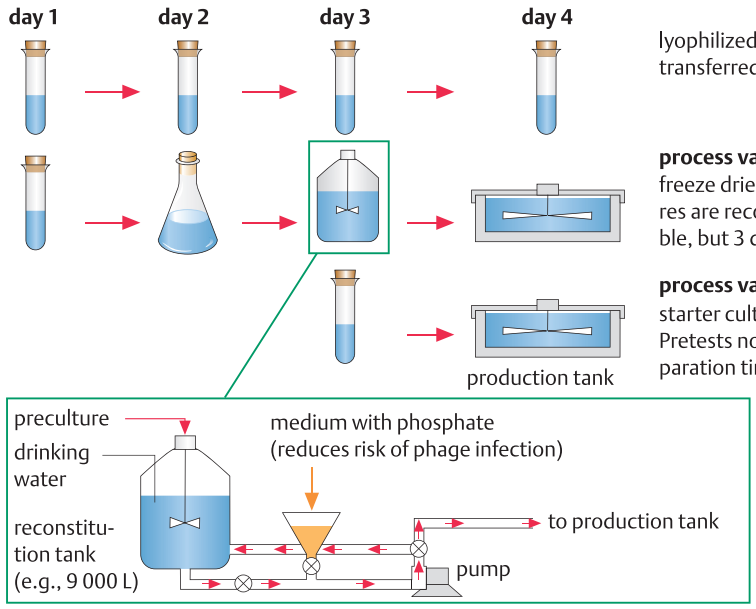
Starter cultures are widely available in the food industry for a broad range of fermented foods. They play a crucial role in the manufacture of fermented milk products (→116) (such as yogurts and cheeses), in the production of sourdough breads (starters) and other baked goods (bakers' yeast) (→120), and in brewing (→112) and wine production (brewers' yeasts) (→110). Starter cultures can be classified into single-strain, single-species, and mixed-species cultures. The key requirements for a starter culture are: a fast and reliable start to the fermentation, reliable manufacture of the desired products, and resistance to antibiotics and phages. The global production value of starter cultures exceeds 1 billion US-$. The legal approval of microorganisms as a starter culture is regulated by an international framework based on pragmatic considerations (e. g., "strain has traditionally been used in fermented foods"). According to an inventory of the *International Dairy Federation,* presently 195 bacterial and 69 yeast and fungal species are classified as safe.

Sausages. The types of sausages that can be stored without refrigeration are mostly produced by adding starter cultures of Staphylococci (e. g., *Staphylococcus carnosus*), although Lactobacilli and *Penicillium* strains are also used. Lactic acid is produced through fermentation from glycogen stored in muscle tissue and lowers the pH to < 5; as a consequence, contamination with acid-sensitive microorganisms is prevented, and muscle protein (isoelectric point: pH 5.3) is gelatinized. Metabolites of the fermentative degradation of fats and proteins play an important role in the flavor of sausages. Starter cultures of salt-stable Staphylococci and Lactobacilli are in use for the production of salted meat and sausages (preservation by the addition of salt, nitrate, or nitrite).

Cheese. In 2010, ca. 14.3 million t of cheese were produced, ca. 6.8 million t in the EU countries and ca. 4.6 million t in the USA alone. There are more than 1,000 varieties of European cheeses. The manufacture of cheese is initiated by a spontaneous infection or by the addition of starter cultures to curd (prepared by precipitation from sour milk, via the addition of rennet or recombinant chymosin (→188) and the removal of whey). A great variety of microorganisms are used in these processes, often from *Penicillium* (Camemberts, Roqueforts), *Streptococcus* (Emmentals) or *Lactococcus* (Gouda-type cheese). The traditional crafts vary widely due to differences in the origin of the milk (cow, goat, sheep), the manufacturing process (aerobic, anaerobic, aerobic-anaerobic), and how starter cultures are added (to the surface or by inoculation). Cheese texture is influenced by curd processing; thus, stretching and kneading in hot water leads to a stringy, fibrous body (Mozzarella, Provolone), repeated piling and mixing takes the edges of the cut curd pieces and removes water (Cheddar), and washing in warm water lowers acidity, providing a milder taste (Edam, Gouda).

Non-Western fermentation products. Angkak is a "red rice" produced in China by the inoculation of moist rice with spores of *Monascus purpureus*. It is used as a flavoring, but also as a digestive therapeutic, due to its content of antibiotics. Kishk is an Asian side-dish prepared by fermentation of sprouted wheat by the addition of sour milk. Miso, a Japanese food flavoring, can be prepared by adding *Aspergillus oryzae* to steamed rice. Soy sauce is a highly aromatic protein hydrolysate that has been prepared for over 1,000 years in China (→186). Today, it is manufactured from a mixture of soybeans and wheat. After inoculation with *Aspergillus oryzae* at high humidity at 35 °C, a surface culture forms. It is mixed with an equal volume of salt water (> 13 % salt), and the resulting mash (*moromi*) is fermented for up to a year at room temperature with lactic acid bacteria and yeasts. In Indonesia and Malaysia, the staple food tempeh is manufactured by the fermentation of steamed soybeans and rice with *Rhizopus oligosporus* (→16).

Preparation of starter cultures



lyophilized or liquid cultures are transferred day by day

process variant 1
freeze dried or frozen starter cultures are reconstituted. Pretests possible, but 3 d preparation time

process variant 2
starter cultures are used directly. Pretests not possible, but no preparation time required

Manufacture of cheese

raw milk	starter culture	clotting coagulum	remove whey	salt
pasteurization	ripening	add rennet, cut, stir, stand		form, press, dry, mature

Manufacture of soy sauce

koji	moromi	finishing
soy meal, wheat bran, solid phase fermentation with *Aspergillus oryzae*, 72 h, 35 °C, high humidity	add equal volume of salted water; ripening for a long time	filter, pasteurize, conserve

Fermented foods of non-Western cultures

product	country	use	raw material	microorganism
koji	Japan	starting product for soy sauce, miso	soy meal, wheat bran, toasted rice	*Aspergillus oryzae*, *Aspergillus sojae*
shoyu (soy sauce)	Japan	dark flavorant	koji	*Pediococcus* sp.
miso (soy paste)	Japan	light flavorant	koji	*Aspergillus*, *Lactobacilli*
tofu, sufu	Japan, China	coagulated protein	soy beans, soy milk	*Mucor sufu* and others
natto	Japan	spicy flavorant	toasted soybeans in pine wood leaflets	*Bacillus natto*
tempeh	Indonesia	staple food	cooked soybeans in banana leaves	*Rhizopus oligosporus*
Ang-kak	Indonesia, China	flavorant, colorant, therapeutic agent	toasted rice	*Monascus purpureus*
gari	Nigeria	food trimming	*Manihot utilissima* (cassava)	*Geotrichum*, *Corynebacterium*

Food and lactic acid fermentation

General. In many cultures, the skills to produce fermented milk products or sauerkraut from fermented cabbage, or to enhance the digestibility of beets as an animal feed by fermentation (silage), have been passed on for hundreds of generations. In 1856, Louis Pasteur laid the foundation for a technological exploitation of these traditions when he discovered the lactic acid bacteria. He found that lactic acid fermentation of food products lowers the pH to ca. 4 which protects against infection by most other microorganisms.

Lactic acid bacteria differ in shape, but can be characterized rather well on the basis of their biochemistry and physiology: they are Gram positive and facultative anaerobic bacteria: though they do not contain heme proteins, e. g., catalase, they are able to grow in the presence of O_2. They cleave lactose to glucose and galactose and metabolize these sugars to lactic acid (→148). "Homofermentative" Lactobacilli such as *Streptococcus pyogenes, Lactobacillus casei* and *Lactococcus lactis* form two mole equivalents of lactic acid per mole glucose, "heterofermentative" Lactobacilli such as *Leuconostoc mesenteroides* and *Lactobacillus brevis* form just one mole equivalent. How much L-(+) lactic acid (usually 50–90 %), D-(–) lactic acid or D,L-lactic acid is formed depends on the occurrence of a species-specific lactate racemase. In addition to lactic acid, fermented milk products contain partially hydrolyzed proteins, no lactose, and a benign microflora; they are therefore considered valuable in the human diet.

Fermented milk products. In Europe, the most important products are sour milk and sour cream, yogurt, kefir, and buttermilk (fat content < 1 %). These can be formed by the spontaneous infection of untreated milk. In commercial dairies, they are manufactured from pasteurized milk by adding starter cultures. The ensuing fermentation results in the formation of lactic acid and an acidification to a pH of 4–5. Yogurt products with > 95 % L-(+) lactic acid, produced, for example, by adding *Lactobacillus acidophilus* starter cultures, possibly supplemented with strictly anaerobic *L. bifidus* (which has been found in the intestinal flora of breast-fed babies), is considered to be especially digestible and to stimulate the immune system (→118). Several *Lactobacillus*-fermented drinks are considered beneficial for health via probiotic activity. Another feature of milk fermentation is the formation of flavorings from the action of proteases and lipases contained in the starter culture. Microbial strains of particular importance for this effect are *Streptococci, Lactobacilli, Leuconostoc* and, in some cases, also yeasts.

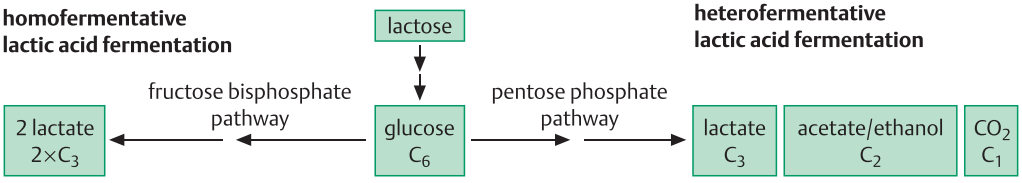
Lactofermented vegetables, fruits, and juices. Important examples are sauerkraut and pickled cucumbers. Sauerkraut is produced in quantities of several million t/y. The manufacturing process usually relies on the spontaneous infection of finely cut stripes of cabbage in large steel tanks or wooden vats (up to 100 t). The microbial flora thus formed are highly diverse and contain lactic acid and other bacteria, yeasts, and molds. The use of starter cultures is being studied. Other examples of lacto- fermented vegetables are: borscht (fermented red-beets; Russia and Poland), Gari (crushed and fermented cassava tubers; Western Africa) and kimchi (fermented Chinese cabbage or radishes; Korea). Lactofermented vegetable juices are storage stable, rich in vitamins and minerals, and well digestible. Examples are tomato and carrot juice.

Sourdough. In contrast to wheat flour, rye flour swells significantly only at pH values < 4.3 – a prerequisite for the formation of elastic, digestible crusts of wheat and rye, cracked grain, and whole-grain breads. For this reason, rye flour is turned into sourdough of pH ca. 4.2 *via* a process based on the combined action of lactic acid bacteria and bakers' yeast.

Silage is a fermented winter feed for cattle. Usually sugar beets are used. They are collected in silos or heaps so as to exclude air, leading to lactic acid fermentation. If the fermentation process is incomplete and lactic acid does not form in sufficient quantities to lower the pH to < 5, Clostridia or yeasts may form, impairing silage quality. Starter cultures are therefore used to improve silage energy contents and stability. Most silages contain the psychotropic pathogen *Listeria monocytogenes*, which can propagate in freezers and thus may lead to the contamination of food products such as soft cheeses, ground meat, and coleslaw, even if they are stored in the cold.

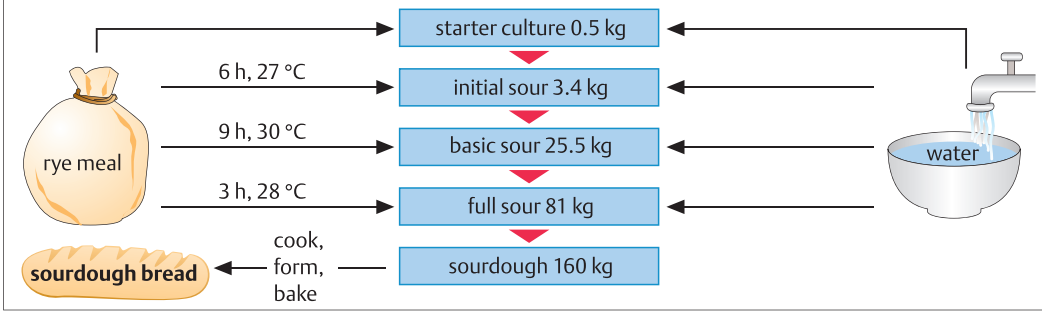
Lactic acid fermentation

homofermentative Lactobacilli		heterofermentative Lactobacilli
Lactococcus lactis *Streptococcus pyogenes* *Lactobacillus delbrueckii* *L. helveticus*	*L. salivarius* *L. casei* and others	*Leuconostoc mesenteroides* *L. lactis* *Lactobacillus brevis* and others



Sourdough

	wheat	rye
% of agricultural area used for production (world)	~ 33	< 3
main product of meal	yeast dough	sourdough
factor responsible for retaining gas on baking	gluten network	pentosanes, proteins formed from acidification
specific volume of bread (volume/weight)	~ 3.5	~ 2.0



Sauerkraut, yogurt and silage

sauerkraut

cut cabbage + 2 – 2.5 % salt → spontaneous inoculation → vat up to 100 t, 18 – 20 °C, protected from air → if lactic acid >1.5 %: filtration, pasteurization, finishing

yoghurt

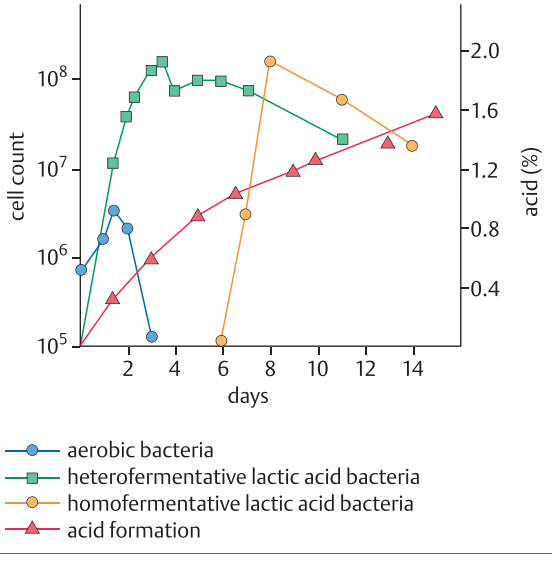
milk
homogenized, pasteurized, with additives, e.g., fruits

batch fermentation	stirred yoghurts
add starter cultures*, fill containers, incubate, cool to 4 °C, distribute	fill tanks, add starter cultures*, incubate, cool to 4 °C, fill containers, distribute

*e.g., *Lactobacillus bulgaricus*, *L. acidophilus*, *Streptococcus thermophilus*

silage

development of microbial population during the fermentation process and acid production



Prebiotics and probiotics

General. Prebiotics are food additives which show a beneficial effect on intestinal flora. Probiotics contain vital microorganisms, e. g., Lactobacilli (→116), which are ingested with food products and show positive health effects. They are thought to affect the large intestine, where a diverse biota of microorganisms ("human intestinal flora" or "human microbiom") (→298) is actively involved in nutrient digestion. The composition of the microbiota has considerable effects on human health and wellness, and is therefore the subject of much contemporary research. The sale of pre- and probiotics is subject to strict legal regulations, in particular with respect to "health claims." Japan is a leader in this area, with over 1,000 approved FOSHU products ("Food of specified health use"), many of which are pre- or probiotics.

Intestinal flora. The microbiom (→298) of human adults consists of about 10^{11} microorganisms. It is composed of some 1,000 different species, with 40–60 dominant genera. The adult human intestinal flora is estimated to weigh around 1 kg, and its cumulative genomes add to some 3 million genes. The intestinal flora of a baby consists predominantly of bacteria from the mother's urogenital tract (in the case of Cesarean delivery, from her skin). When the baby's nutrition changes to mixed food, a different microbiota develops which remains essentially stable during his or her adult lifetime ("robust enterotypes"), though adverse effects such as malnourishment or malnutrition, antibiotics or changing dietary habits during travel may have transitory effects. As an individual ages, their microbiota loses diversity. The enterotpye of an individual is fixed during their first years of life, and is determined by genetic factors and nutritional habits. The intestinal wall remains largely sterile and is protected against bacterial infection by peptide antibiotics ("defensins") (→210). However, bacterial toxins may enter the enterohepatic blood stream of the host together with hydrolyzed nutrients such as amino acids or peptides, if immunoprotection fails or is defeated. Molecular genetic methods, in particular transcriptome analysis and the sequencing of 16S rRNA in metagenomic intestinal DNA libraries (→74), have greatly advanced our knowledge about the composition of the human microbiome in health and disease. Profiling the metabolites in our large intestine (many of which are the product of microbial metabolism) by high-resolution mass spectrometry or NMR (metabolomics, metabonomics) has allowed a rational evaluation of the effects of pre- and probiotics.

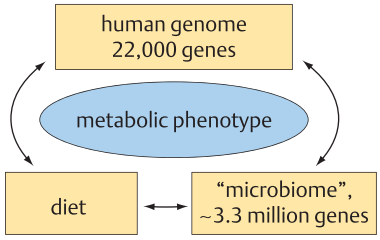
Prebiotics. Typical prebiotics are fructooligosaccharides, inulin, galactooligosaccharides and lactulose. Fructoologisaccharides and inulin occur in plant-derived food such as chicorium, Jerusalem artichoke or Scorzonera. Galactooligosaccharides and lactulose are industrially produced from lactose. They act as selective growth promoters for "good" intestinal bacteria such as Bifidobacteria and Lactobacilli.

Probiotics contain predominantly vital Bifidobacteria and Lactobacill and are mostly consumed as fermented food products or drinks, occasionally also as food additives. As most microorganisms are inactivated during passage through the acidic stomach, their effect on the large intestine is debated. Acid-stable probiotic strains are also being used in some commercial products.

Regulations. Since July 2007, so-called *health claims* are strictly regulated in the European Union ("Regulation on Health Claims and Nutrition"). Health-related claims on food or food components (including probiotics) are now only permitted if they have been scientifically proven. At present, all health-related probiotics claims are being compiled and summarized. This list is submitted to EFSA (European Food and Safety Authority) (→334) which must judge the scientific soundness of each claim. For some probiotics, EFSA has already published negative decisions.

FOSHU ("food for specified health use") is an official status for Japanese food products or additives which have been proven to maintain or improve human health. To obtain FOSHU status for a product, extensive data must be submitted for evaluation to the governmental Consumers Affairs Agency and, if a positive decision is reached, acknowledged by the Ministry of Health, Labor and Welfare. The 1,000 FOSHU products in Japan include not only "foods to modify gastrointestinal conditions" (pre- and probiotics), but also products which enhance human health in other ways, e. g., bone formation through osteogenesis, or blood pressure regulators.

Man, microbiome and diet

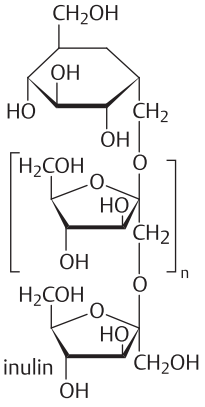


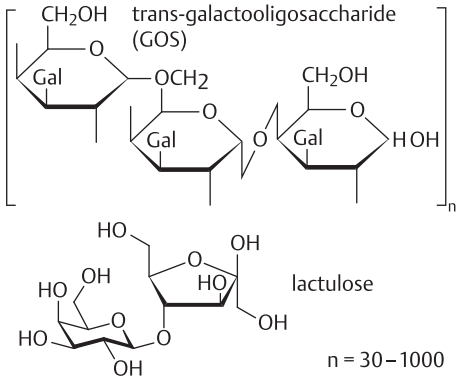
Probiotic bacteria (examples)

Probiotic bacteria as health food drinks	
Bifidobacterium animalis subsp. lactis HN019	Danisco
Lactobacillus acidophilus LA5	Chr. Hansen
Lactobacillus johnsonii La1 (Lactobacillus LC1)	Nestlé
Lactobacillus casei DN-114001 (= *Lactobacillus defensis*, Actimel)	Danone
Lactobacillus casei Shirota (DSM 20312)	Yakult

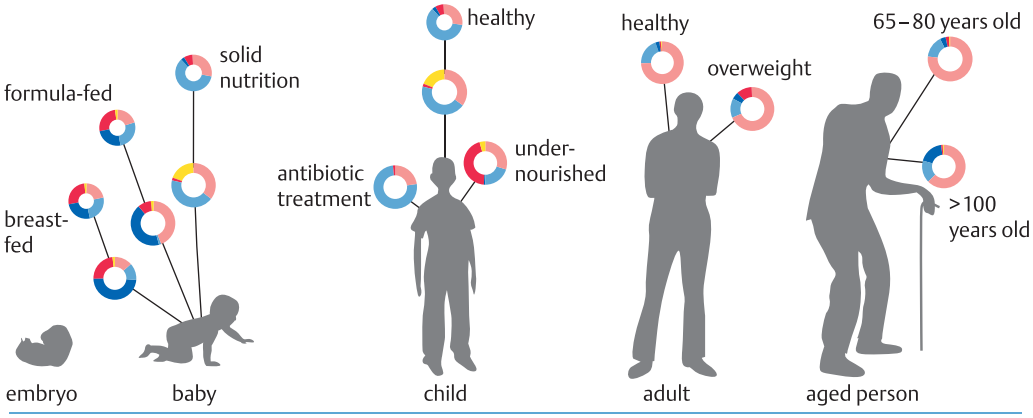
Prebiotic oligosaccharides and sugars (selection)

Prebiotic oligosaccharides (selection)
inulin
trans-galactooligo-saccharide (GOS) fructo- (FOS), xylo-(XOS), mannan-oligosaccharide (MOS), polydextrose
disaccharides such as lactulose, monosaccharides such as tagatose





Development of the intestinal flora during life



- Firmicutes (mostly Gram-positive)
  Clostridium, Bacillus, Lactobacillus, Staphylococcus, Streptococcus, Pediococcus
- Bacteroides (Gram-negative)
  Bacterioides fragilis
- Actinobacteria (Gram-positive)
  Bifidobacteria, Actinomycetes
- Proteobacteria (Gram-negative)
  Escherichia, Salmonella, Vibrio, Heliobacter
- others

FOSHU Health claims

selection	functional components
related to gastrointestinal tract	oligosaccharides, lactose, Bifido- and Lactobacteria, difficult-to-digest dextrins, guaran
related to blood sugar	difficult-to-digest dextrins, wheat albumen, tea polyphenols, L-arabinose
related to blood pressure	lactotripeptide, casein peptides
related to dental hygiene	palatinose, maltotriose, erythritol

Bakers' yeast and fodder yeasts

General. The preparation of dough from flour is documented back to early history: Egyptian clay tablets tell about the baking of "beer breads", where moist barley exposed to a yeast fermentation was used. Until the 19th century, bakers in Europe also used brewers' yeast for baking. It was obtained either by the "Dutch process" (ca. 1750), based on filtration of a turbid mash, or by the "Vienna process" (ca. 1800), where yeast was skimmed off the fermentation vat. Around 1870 Louis Pasteur in France discovered that the production of baker's yeast (*Saccharomyces cerevisiae*) (→14) in high yield with little concomitant formation of ethanol was possible with strong aeration. Since then, baker's yeast has been produced in stirred vats under aeration. Molasses (→88) is widely used as a carbon source. Substrains of *Saccharomyces cerevisiae* are industrially manufactured as baker's yeast, brewer's yeast, as starter cultures for the production of bio-ethanol and as a nutritional additive to health products and animal feed. The global yeast market is in the range 4 billion US-$ (2013), with baker's yeast constituting about 1/3 of the overall value. The first processes for preparing yeasts as animal feeds ("fodder yeasts") were developed in the 1930s for the purpose of creating a self-sufficient economy. In the post-war period, a new driving force appeared: closure of the "protein gap" between developed and developing nations. More recently, the sustainable use of biomass residues has become a major driver. Thus, yeasts grown by solid-state fermentation on agricultural residues in biorefineries are being explored as animal feed or fertilizer.

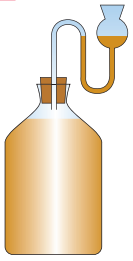
Bakers' yeast fermentation. Due to its favorable price and high content of sugar, nitrogen sources, vitamins, and minerals, molasses (from cane or beet sugar) has become an important raw material for yeast production. It is produced by extraction of chopped sugar cane or sugar beets with hot water and contains 40–50 % saccharose, which is hydrolyzed by yeast invertase into glucose and fructose that can be metabolized by yeast.

Technology. To obtain a high yield coefficient (Y_S) (→92) for the transformation of molasses into yeast cell mass, two factors are essential: strong aeration during fermentation to suppress ethanol formation, and strict control of the glucose concentration in the medium. If glucose concentrations exceed 100 mg L^{-1}, respiration is decreased even at high oxygen concentrations, and ethanol is formed (catabolite repression, "aerobic fermentation," "Crabtree effect") (→92). As a result, modern fermentation procedures use optimized aeration and stirrer systems (→94) and a computer-based feed control for the molasses. In an optimized process, 1 kg H_{27} of yeast per kg molasses is formed. H_{27} is a relative value used in the yeast industry to calculate yields, indicating pressed yeast with 27 % dry weight. Since molasses contains ca. 50 % sugar and since the yield of yeast relative to sugar substrate is ca. 54 g dry weight/100 g sugar, the yield of yeast is ca. 1 kg H_{27} kg^{-1} molasses. Traditionally, workup procedures led to a pressed yeast having limited storage stability, necessitating local production sites for distribution. Today, the cell mass is dried in a vortex dryer to yield dried baker's yeast that is stable for several months and is reactivated by water and sugar within a few minutes, an important convenience feature for bakers who begin working early in the morning.

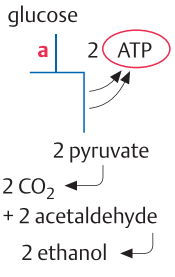
Fodder yeasts compete in price with other protein-rich residual materials such as soy meal. As a result, the price of the carbon source is critical for the competitiveness of a fodder yeast manufacturing process. Oil fractions, natural gas, ethanol, methanol, celluloses, hemicelluloses, starch, and whey were studied as carbon sources. In the Swedish Symba process, the yeast *Endomycopsis fibuliger* degrades potato starch into glucose on which *Candida utilis* (→144) grows and accumulates cell mass. Another carbon source is spent sulfite liquor used in in pulp manufacturing from wood chips, which contain ca. 4 % polyoses. In the Finnish *Pekilo* process, the fungus *Paecilomyces variotii* is used for this purpose. In the Canadian *Waterloo* process, the fungus *Chaetomium cellulolyticum* is used to produce cell mass from agricultural and forest waste materials such as straw, bagasses, manure, or sawmill chips. The Rank-Hovis McDougall process is based on the fungus *Fusarium venenatum*, which transforms glucose to the low-calorie health food *Quorn*® which offers a meat-like texture.

Baker's yeast

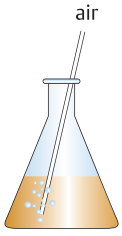
A fermentation



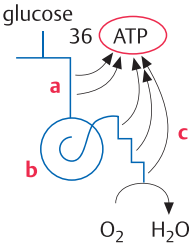
anaerobic yeast



B respiration



aerobic yeast

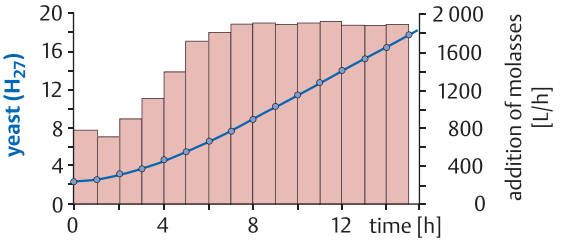


a glycolysis
b citric acid cycle
c respiration chain

procedure	yeast H_{27}* [kg]	ethanol [L]
A Dutch procedure (until ~ 1850)	3 – 6	25 – 30
Vienna procedure (until ~ 1915)	14	25
Copenhagen procedure (from 1877)	20	20
B modern fed-batch procedures	up to 85	< 1

* in modern procedures, 27 kg yeast dry mass (H_{27} yeast) is formed from 100 kg of molasses

yeast manufacture with computer-assisted addition of molasses



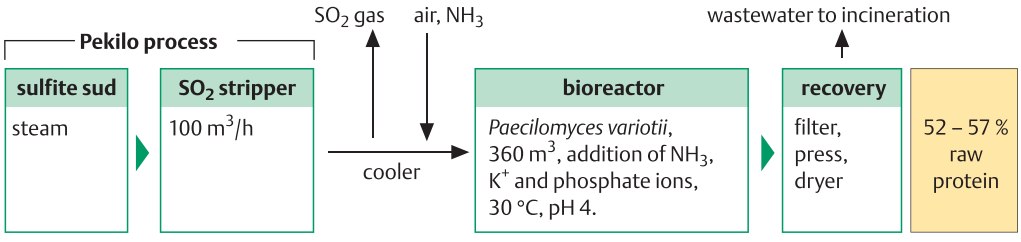
preculture	bioreactor	recovery	by-products
"mother yeast"	> 100 m^3, molasses, N-source	separator	ethanol, wastewater, draff

pressed yeast H_{27}	dried baker's yeast	fodder yeast
from vacuum rotary drum filter	after extrusion, vortex, dryer, cooler	after pasteurization

Fodder yeasts

raw material	yeasts and fungi	remarks
potato starch	*Endomycopsis fibuliger* and *Candida utilis*	**Symba process**, Sweden. *E. fibuliger* degrades starch, *C. utilis* grows faster and forms cell mass from glucose
whey	*Kluyveromyces fragilis*	for cleaning wastewater from dairies. Product contains yeast and whey residues
spent sulfite liquor	*Saccharomyces cerevisiae*, *Candida utilis*, *Paecilomyces variotii*	hexose residues of wood hydrolysis are metabolized by *S. cerevisiae*, pentoses are metabolized by *C. utilis* **Pekilo process**, Finland: continuous process to manufacture fodder yeasts from sulfite pulping waste
purified glucose	*Fusarium graminearum*	**Rank Hovis McDougall process**, United Kingdom: use of the fibrous texture of fungal mycelium as "functional protein"

Pekilo process



reduction of biologically degradable components (BOD_5) of waste liquor by 85 %

Fodder yeasts from petroleum feedstocks, single cell oil

General. After World War II, technological advances in petrochemistry and fermentation technology led to the idea of developing single cell protein (SCP) by fermentation. This product would help to close the gap in protein supply between affluent and poor populaces in a world where the human population showed exponential growth. The potential of single cell oil (SCO) had been explored during the war in Germany, in answer to the Allies' sea blockade which had halted plant-oil imports from Asia and the Americas. The search for renewable energy resources has revitalized this field.

Single cell protein: raw materials and microorganisms. Yeasts such as *Candida tropicalis* and *C. bombicola* (→14) can grow on alkanes, which are terminally oxidized within the peroxisomes to mono- and dicarboxylic acids and then metabolized further. Based on these observations, the use of high-boiling petroleum fractions as a C-source for yeast protein formation was intensely explored between 1965 and 1975. As a result of the two oil crises, petroleum was gradually replaced by methanol. This C-source can be assimilated by a wide range of methylotrophic bacteria (→14) and yeasts such as *Hansenula polymorpha, Pichia pastoris, Candida boidinii, Methylophilus methylotrophus*, and *Methylomonas clara*. The metabolism of methanol starts with oxidation to formaldehyde, which is then incorporated into the pentose phosphate pathway. Formaldehyde can also be further oxidized to CO_2 via formic acid, providing reduction equivalents.

Yeast cell mass from alkanes. High-boiling alkanes are only sparingly soluble in water, but emulsifiers should not be added since they might contaminate the product and may also lead to foam problems during fermentation. As a result, optimization of aeration and of the mechanical stirrer system is of key importance in process development. Since the aeration level in the process is high (10 vvm, due to the requirement of 16 mol O_2 per mol hexadecane), formation of foam must be carefully controlled. The 40 m^3 bioreactors in Grangemouth, UK (BP) and a 100 m^3 bioreactor plant in Sardinia used Intermig stirrers and mechanical foam separators. With fed-batch cultures fermented for 5 d, 0.9 kg of moist *Candida tropicalis* kg^{-1} alkane was produced. Cell mass was recovered using separators. Since the high RNA content (4%) of the product was critical from a dietary point of view, it was reduced by limited autolysis. The yeast was stabilized for transport by drying.

Methanol-based processes. Methanol has a low boiling point (64.5 °C) and is toxic even for methylotrophic microorganisms in concentrations > 100 mg L^{-1}. Using a narrow, tall airlift fermenter (ICI Billingham, UK: 8 × 60 m) and up to 600 computer-controlled nozzles, the distribution of methanol was controlled, and its solubility was enhanced due to the high hydrostatic pressure. The considerable aeration rate required was achieved by a loop system. Using the bacterium *Methylophilus methylotrophus*, a continuous 2-d process resulted in 0.5 kg biomass kg^{-1} methanol. Cell mass was separated using separators, the RNA content was reduced by limited autolysis, and a vortex dryer was used to render the protein product stable for transport.

Acceptance and economics. Alkane yeasts were criticized early as a food, in view of a potential enrichment of carcinogenic polyaromatic compounds from petroleum. Though in-depth toxicological studies did not support this view, registration in 1974 was limited to use as an additive in domestic animal feeds. The price increases of petroleum during the two oil crises finally led to termination of the projects. The market introduction of methanol-derived biomass did not meet with such difficulties, as it was declared a feed additive from its beginning. However, rising methanol prices and an EU-wide subvention of milk powder as a feed protein made this process also uneconomic.

Single cell oil can be obtained from glucose-using yeasts such as *Rhodotorula glutinis* or fungi such as *Mortierella isabellina*, as well as from CO_2 using cyanobacteria (→18). Oil yields > 60% of dry cell mass have been reported, and time-volume yields have been increased by high-cell density culture, e. g., in view of biodiesel production. The triglyceride composition (→62) can be modulated by genetic engineering. Thus, *M. isabellina* has been engineered to produce high amounts of ω-3 fatty acids, which are components of health food.

Single cell protein (SCP)

carbon-source	microorganisms	problems
high-boiling alkanes	*Candida lipolytica, Candida tropicalis*	isolation difficult (formation of emulsions), oil residues, taste, consumer acceptance
methane	*Methylococcus capsulatus*	high O_2 consumption, strongly exothermic process (cooling), danger of explosions
methanol	*Hansenula polymorpha, Pichia pastoris, Candida boidinii, Methylophilus methylotrophus*	high content of RNA

Composition (%) of soy meal and milk powder as compared to single cell protein

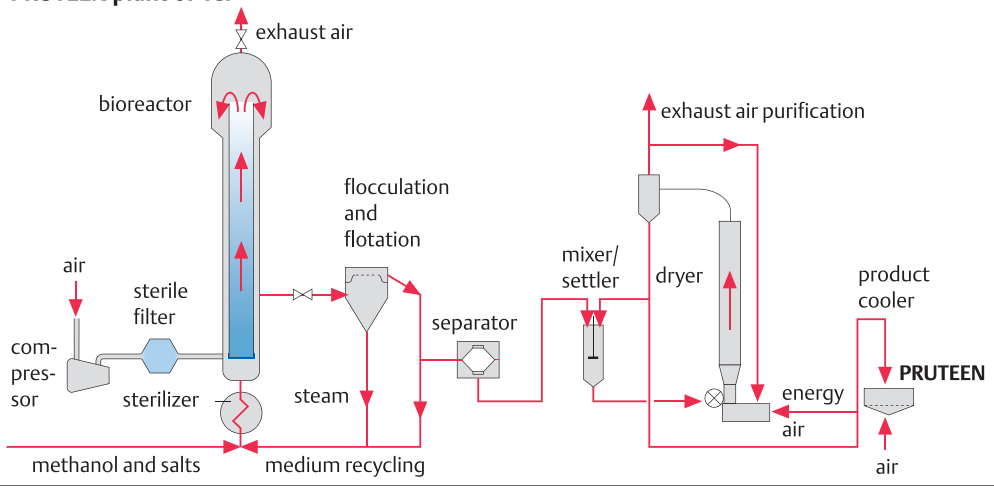
SCP-product	raw protein	amino acids	fats	nucleic acids	salts, water
TOPRINA (BP)	60	54	9	5	10
PRUTEEN (ICI)	83	65	7.4	15	10
soy meal	45	40	2	-	18
whole milk powder	25	-	27	-	10

SCP yields (pilot and production plants)

microorganism	carbon source, metabolism	Y_s *	productivity [kg/m^3 · h]
Candida lipolytica (TOPRINA)	alkanes	0.95	2
Methylophilus methylotrophus (PRUTEEN)	methanol, RMP	0.53	8 – 10
Candida utilis	ethanol	0.8	4.5
Saccharomyces cerevisiae (bakers' yeast)	molasses, FDP	0.85	5.2
Paecilomyces variotii (PEKILO)	sulfite suds, FDP	0.6	2.7

*g cell dry weight/g carbon source
RMP: ribulose monophosphate pathway, FDP: fructose diphosphate pathway

PRUTEEN plant of ICI



Manufacture of PRUTEEN

raw materials	airlift reactor	tion and flotation	isolation	
methanol from 600 nozzles, ammonia, salts	3,000 m^3 working-volume; 60 m high, 600 t	30 min at 64 °C, RNA content is reduced from 80 to 2 mg/g	separation, drying	83 % raw protein, < 1 % nucleic acids

Amino acids

General. Amino acids have been used for medical purposes, e. g., in infusion preparations, ever since their important metabolic role was discovered in the first half of the 20th century. Some amino acids such as D,L-methionine, L-lysine, or L-threonine serve as additives in animal feed. The findings that L-glutamate has taste-enhancing properties in food and that the dipeptide Aspartame™ is an excellent low-calorie sweetener helped promote the industrial production of amino acids. The 20 proteinogenic amino acids are the building blocks of proteins and peptides (→28). Most higher organisms depend on the uptake of several of these amino acids with their food (essential amino acids). In man and in many of his livestock, these are L-methionine; L-lysine; the aromatic amino acids L-phenylalanine, L-tyrosine, and L-tryptophan; and the hydrophobic amino acids L-valine, L-leucine, and L-isoleucine. An "amino index" has been developed which shows the limiting amino acid in human nutrition and can be monitored experimentally. Non-proteinogenic amino acids having the D-configuration at their Cα atom, occur in natural compounds. They serve as chiral synthons in organic synthesis, e. g., for the manufacture of semi-synthetic antibiotics (→132).

Economic considerations. The annual production of amino acids is ca. 5 mill. t/y (2012), and their market volume approaches 10 bill. US$. Most commercial producers are located in Northern Asia. The most important product is sodium-L-glutamate (ca 2.5 mill. t) (→126), followed by L-lysine (ca 1.5 mill. t) (→128) and D,L-methionine (ca. 900,000 t). L-aspartic acid and L-phenylalanine, starting products for the production of the sweetener Aspartame™, are produced at 15,000 t each (→130). Only a small proportion of these amino acids is prepared pyrogen-free and at a high purity for medical-therapeutic applications, e. g., for infusions. The largest and still growing share of industrially produced amino acids is used as an additive in animal feed (lysine, methionine, threonine, tryptophane) or is used for human nutrition (glutamate, Aspartame).

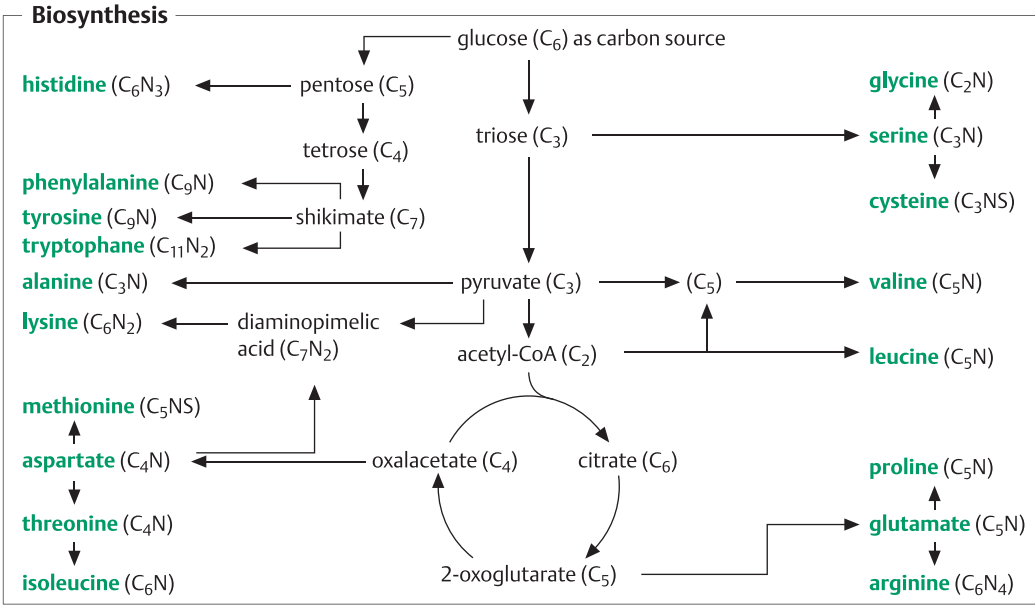
Production. Four different methods are available for manufacturing amino acids: 1) extraction from protein hydrolysates, 2) chemical synthesis, 3) biotransformation of chemical precursors using enzyme or cell reactors, and 4) microbial production via fermentation. The extraction of amino acids from protein hydrolysates (the chiral pool) was used especially for L-cysteine and L-cystine, L-leucine, L-asparagine, L-arginine, and L-tyrosine. Various plant proteins and proteinaceous wastes from the slaughterhouse were used as starting material. After acid hydrolysis, the hydrophobic amino acids L-phenylalanine, L-leucine, and L-isoleucine were separated first by precipitation and ethanol extraction. The water-soluble amino acids were separated in a basic, an acidic, and a neutral fraction by ion-exchange chromatography, followed by crystallization. Chemical synthesis usually leads to racemic amino acids. They can be commercially used if the lack of chirality does not impede function, as is true with D,L-alanine (for the taste finish of fruit juices), but in particular for D,L-methionine as a feed additive. Racemic amino acids are separated into the two enantiomers differing at the C_α-atom by using a biocatalyst. The biocatalyst used is either an isolated enzyme or whole cells containing an appropriate enzyme. Under commercial conditions, it is usually preferable to immobilize the biocatalyst, since this allows for a continuous process and a high operational lifetime of the catalyst. The economic success of such processes is usually based on the simple and inexpensive synthesis of the chemical precursor. Due to advances in recombinant technologies, nearly all proteinogenic amino acids are now manufactured via fermentation processes as the most economic choice. *E. coli* and *Corynebacterium glutamicum* (→20) are the preferred host organisms, and producer strains have been optimized through metabolic engineering. The only exceptions are glycine, D,L-methionine and L-aspartate which are produced by chemical synthesis. The genome sequence analysis of *Corynebacterium glutamicum*, the industrial strain for the production of glutamic acid and lysine, was completed in 2003 by several competing groups, and this achievement has led to production strains with even higher yields. Since the pathways involved in amino acid biosynthesis and secretion are well understood, metabolic engineering (→318) is widely used to direct metabolic fluxes into the direction of the desired amino acid.

Industrially important amino acids

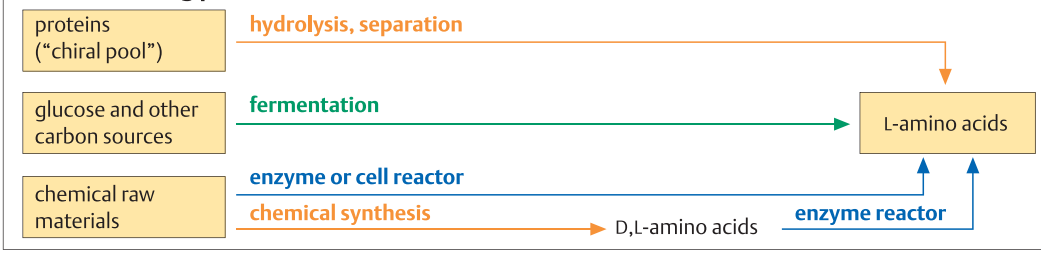
amino acid	annual production* [t/y]	value* [US$/kg]	manufacturing procedure	main application
proteinogenic amino acids				
L-glutamate	2,500,000	1	fermentation	taste enhancer
L-lysine	1,300,000	2	fermentation	feed additive
D,L-methionine	900,000	2	chemical synthesis	feed additive
L-threonine	15,000	5	fermentation	feed additive
L-aspartic acid	15,000	10	chiral pool, cell reactor	Aspartame™
L-phenylalanine	15,000	10	fermentation	Aspartame™, medicine
glycine	16,000	10	chemical synthesis	sweetener
L-arginine	2,000	20	fermentation	medicine, cosmetics
L-tryptophan	4,500	20	fermentation	feed additive
all others	~10,000		chiral pool, fermentation, enzyme and cell reactors	medicine and other applications
other amino acids (examples) D-phenylglycine, D-4-hydroxyphenylglycine			chemical synthesis, enzyme reactor	precursor for Ampicillin, Amoxicillin
(*S*)-5-hydroxytryptophan			chemical synthesis, plant extract	Oxitriptan™, an anti-depressant

*estimates, 2012

Biosynthesis



Manufacturing processes



L-Glutamic acid

General. L-glutamic acid is a proteinogenic amino acid. In higher organisms, one of its key function is to act as a neurotransmitter in the brain. The brain is protected from the influx of excess L-glutamic acid by the blood–brain barrier (BBB) and forms L-glutamic acid from L-glutamine, which can cross the BBB. As early as 1908, the flavor-enhancing effect ("umami") of *Konbu* algae was discovered in Japan and traced to the presence of sodium glutamate. In 1909, Ajinomoto started to produce this amino acid from acid hydrolysates of wheat gluten and soy protein. In 1957, researchers at Kyowa Hakko discovered that *Corynebacterium glutamicum* (→20) secretes L-glutamic acid when grown on sugar-containing media. In subsequent decades, this strain was improved by mutagenesis, and the fermentation technology was optimized, resulting in yields of sodium glutamate up to 150 g L^{-1}.

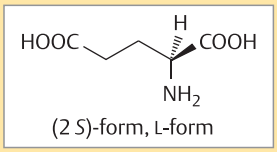
Organism and biosynthesis. *C. glutamicum* forms glutamic acid as a by- product of the citric acid cycle, via isocitrate and 2-oxoglutarate. In wild-type strains, the ratio of glutamate biosynthesis and oxidation of C_2 units via the citric acid cycle is strictly regulated. Industrial mutants, in contrast, exhibit the following properties: 1) they secrete glutamate much better, 2) key enzymes of the biosynthetic pathway are deregulated, and 3) anaplerotic pathways (fill-up reactions) are activated. Further details regarding these include (respectively): 1) mutants with enhanced membrane permeability have been obtained using various measures such as reducing the availability of biotin, oleic acid, or glycerol (using oleic acid- or glycerol-auxotrophic strains), using strains with deformed cell walls, adding penicillin; 2) in production strains, the activity of 2-oxoglutarate dehydrogenase is much lower than that of L-glutamate dehydrogenase (K_m ca. 70 fold, v_{max} ca. 150 fold); 3) carboxylation of phosphoenol pyruvate (PEP) via PEP carboxylase and activation of the glyoxylate cycle lead to enhanced formation of oxaloacetic acid, the precursor of citric acid, from glycolysis. Since PEP carboxylase requires biotin, this cofactor must be available in sufficient quantities. In addition, several of the enzymes involved in glutamate biosynthesis have been deregulated towards intermediary metabolites, various end products, NH_4^+, and the NAD^+/NADH pool. Since the genome of *C. glutamicum* (3.1 million bp) has been completely sequenced in 2003, these classical methods were increasingly being complemented by genetic and metabolic engineering. Thus, the effect of multi-copy gene cassettes of glutamate dehydrogenase on glutamate productivity was investigated. Moreover, phosphorylation of the regulatory protein OdhI, as well as mutants of the mechanosensitive export channel YggB, were discovered to influence glutmate production and export, respectively. Discoveries such as these enhance functional understanding of the biosynthetic pathway and are continuously being used for the generation of new high-performance industrial strains.

Fermentation and recovery. Molasses or starch hydrolysates are often used as a carbon source (→88). Under optimized culture conditions, high-performance mutants of *C. glutamicum* convert 60–70 % of these carbon sources to glutamate. Ammonia is used as a nitrogen source. The biotin content of the medium is optimized, and the pH value is kept at 7.8. The oxygen supply is critical (optimal k_d value 3.5×10^{-6} mole O_2 atm^{-1} min^{-1} mL^{-1}). Production is carried out in bioreactors of up to 500 m^3. Preculture fermenters of increasing volume are used for inoculation. To prevent catabolite repression, a fed-batch mode (→92) is preferred for the process, and after ca.14 h (after formation of the cell mass) the glucose content of the medium is kept at 0.5 %, by monitoring the CO_2 content of the exhaust gas. Glutamic acid, dissolved in the fermentation as ammonium glutamate, is isolated from the medium (ca. 150 g L^{-1} after 60 h) by separation of cells via ultrafiltration, followed by anion-exchange chromatography and crystallization. The anion-exchanger is recycled to its sodium form, recovering the ammonia.

Economic considerations. Monosodium glutamate (MSG) is mainly used as a flavor enhancer in the food industry, often in combination with nucleosides such as IMP or GMP. Commercial production of L-glutamic acid in 2009 was around 2.3 mill. t, with China being the largest producer (ca. 700,000 t). Most other producers are also located in Asian countries, where MSG is mostly consumed. At a price of ca. 1,000 US$ t^{-1}, the market volume exceeds 2.3 billion US$.

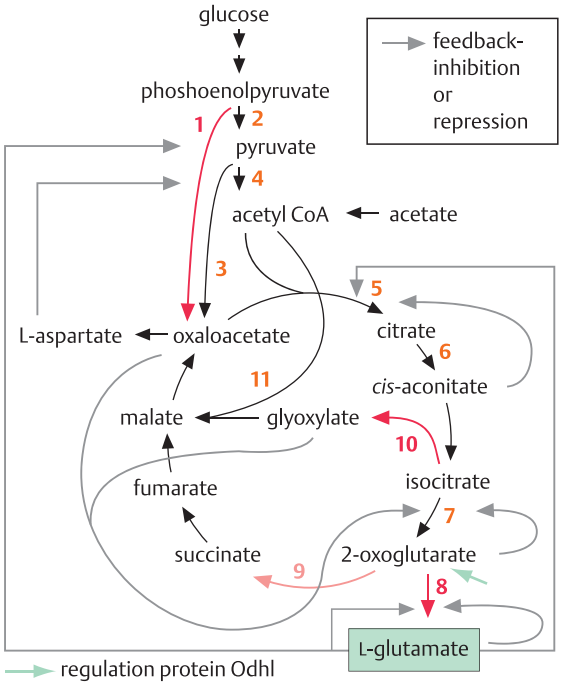
L-Glutamic acid

$C_5H_9NO_4$			
M_R	147.13	CAS	56-86-0 (L)-form
m. p.	247 – 249 °C (decomp.)		
solubility	600 g/L water (20 °C)		



(2 *S*)-form, L-form

Biosynthesis and high-performance mutants

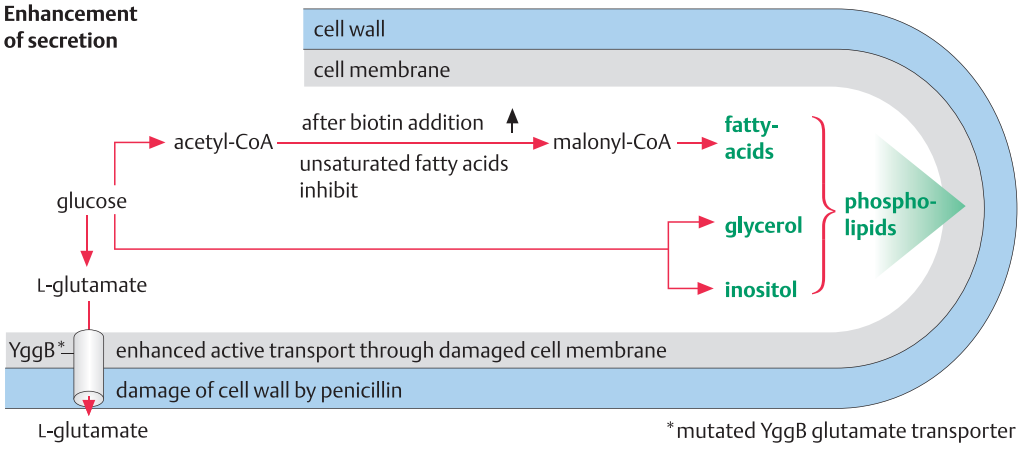


enzyme	gene
1 phosphoenolpyruvate carboxylase (PEPC)	*ppc*
2 pyruvate kinase	*pyk*
3 pyruvate carboxylase	*pyc*
4 pyruvate dehydrogenase	*pdh*
5 citrate synthase	*gltA*
6 aconitase	*citB*
7 isocitrate dehydrogenase	*icd*
8 L-glutamate dehydrogenase (GDH)	*gdh*
9 α-ketoglutarate dehydrogenase (KDH)	*aceE*
10 isocitrate lyase (ICL)	*aceA*
11 malate-synthetase (MS)	*aceB*

high-performance mutants:

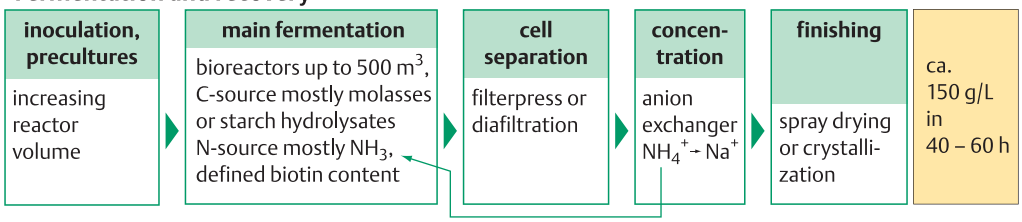
1. higher activities of **PEPC**, **GDH**, **ICL** and **MS**
2. reduced activities or block at **KDH**
3. **PEPC** with reduced feedback inhibition by L-glutamate

Enhancement of secretion



*mutated YggB glutamate transporter

Fermentation and recovery



D,L-Methionine, L-lysine, and L-threonine

General. These three amino acids are mainly used as feed additives. For humans and most domesticated animals, these are essential amino acids, which are not produced by these organisms. Many crops used as food or feeds, e. g., corn, soybeans, oats, barley, wheat, and rice, do not contain enough of these amino acids for healthy nutrition. It is thus recommended that predominantly vegetarian diets be supplemented with these amino acids. In feeds, this deficiency plays an even more important role, since the increased mass during fattening of an animal on wheat or rice reaches the nutritional standard of casein only if L-lysine and L-threonine are added. Similarly, a corn-based feed requires the addition of D,L-methionine, L-lysine, and L-tryptophan. These amino acids are produced industrially by chemical synthesis or fermentation.

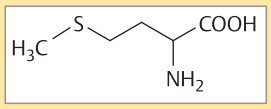
D,L-Methionine. The synthesis consists of five steps involving acrolein, methanethiol, and HCN, leading to the intermediary formation of a hydantoin. Since D-methionine is converted to L-methionine by higher animals, racemic D,L-methionine can be used as a feed additive; separation of the enantiomers is not necessary. Small amounts of L-methionine are manufactured by hydrolyzing derivatives of D,L-methionine by enzymes which selectively cleave the L-enantiomer derivative.

L-Lysine is mainly produced by fermentation using mutants of *Corynebacterium glutamicum* (→20). It is formed via oxaloacetate of the citric acid cycle through condensation of aspartic and pyruvic acid in a multi-step pathway leading through diamino pimelic acid (DAP) as an intermediate. Branches of this pathway also lead to L-threonine and L-methionine, which can suppress the formation of lysine by feedback inhibition. In lysine-overproducing mutant strains, this regulatory process is eliminated due to deregulation or by bypassing strongly regulated enzymes via auxotrophic block mutants. S-Aminoethyl cysteine (AEC)-resistant mutants play a special role: this antimetabolite binds as L-lysine to the allosteric center of aspartate kinase. Consequently, AEC-resistant mutants are no longer inhibited by L-lysine, resulting in higher lysine yields. Since the genome sequence of *C. glutamicum* is known and annotated as of 2003, and all genes coding for the enzymes involved in this path- way have been cloned, genetic and metabolic engineering methods based on flux analysis play a most important role in achieving even more potent mutant strains. Today, yields are in the range of > 100 $g\,L^{-1}$ in 60 h, and (as for L-glutamate) repeated fed-batch protocols in bioreactors up to 750 m^3 in volume are used. The carbon source is usually sugarcane molasses, and substrate conversion may reach 75 g L-lysine per 100 g glucose. The content of biotin in the medium must exceed 30 $\mu g\,L^{-1}$. Recovery may proceed by spray-drying and granulation of the whole broth, or by separation of cells followed by ion-exchange chromatography and crystallization. An interesting alternative, though no longer economically viable, is the production of lysine from D,L-α-amino-ε-caprolactam (ACL), an inexpensive intermediate from the chemical synthesis of Nylon™, with acetone-dried cells of *Cryptococcus laurentii* in a cell reactor. After enantioselective hydrolysis, re-racemization of the remaining D-α-amino-ε-caprolactam with a D-aminocaprolactame racemase, using cells of *Achromobacter obae,* leads to nearly quantitative yields of L-lysine.

L-Threonine. The organisms of choice for producing this amino acid are deregulated mutants of *Escherichia coli.* The best strains form > 100 $g\,L^{-1}$ in only 30 h. The threonine operon has been cloned and is used for further strain improvement. Recovery is initiated by cell separation, followed by ultrafiltration and then crystallization of the product from the ultrafiltrate.

Economic considerations. Annual production of L-lysine in 2009 was ca. 1.3 mill. t, of D,L-methionine ca. 850,000 t, and of L-threonine ca. 190,000 t. The preferred process for methionine production is chemical synthesis of the racemate. L-Threonine and L-lysine are produced exclusively by fermentation. The price of the three amino acids is between 1,500 and 4,000 US$ t^{-1}, leading to a market value of about 3 billion US$. In the more distant future, the addition of industrially produced amino acids to food and feeds is expected to receive competition from the generation of transgenic crops containing an amino acid composition optimized for nutritional purposes.

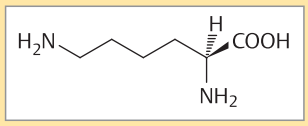
D,L-Methionine



$C_5H_{11}NO_2S$

m.w.	149.21
m.p.	280 – 281 °C, dec.
CAS	63-68-3

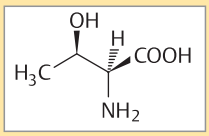
L-Lysine



$C_6H_{14}N_2O_2$

m.w.	146.19
m.p.	224 °C, dec.
CAS	56-87-1

L-Threonine



$C_4H_9NO_3$

m.w.	119.12
m.p.	255 – 253 °C, dec.
CAS	72-19-5

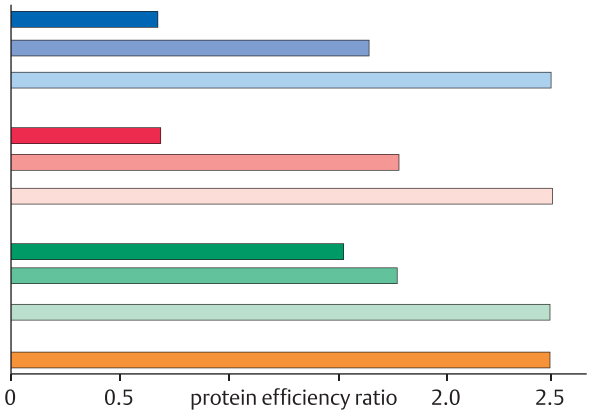
Amino acids in animal feed

wheat without additions
addition of 0.25 % *L*-lysine
addition of 0.4 % *L*-lysine
and 0.4 % *L*-threonine

corn without additions
addition of 0.05 % *L*-lysine
addition of 0.4 % *L*-lysine
and 0.035 % *L*-tryptophan

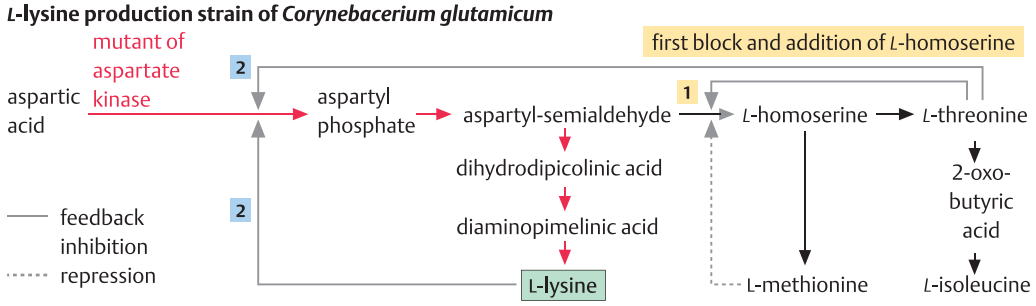
rice without additions
addition of 0.1 % *L*-lysine
addition of 0.2 % *L*-lysine
and 0.1 % *L*-threonine

for comparison: **casein** without any additives



Deregulated biosynthesis

L*-lysine production strain of *Corynebacerium glutamicum



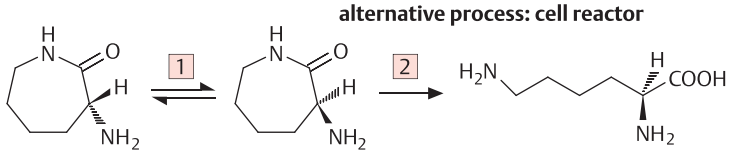
mutants where *L*-lysine biosynthesis is no longer repressed by AEC, *L*-lysine and *L*-threonine, obtained through screening of mutants in the presence of S-aminoethyl cystein (AEC), an antimetabolite

Fermentation and recovery of *L*-lysine

preferred process: fermentation

production strain	30 m³ bioreactor	750 m³ bioreactor	cell separation	recovery	
auxotrophic, regulatory and transgenic mutant of *Corynebacterium glutamicum*	dextrose, soybean meal	dextrose, soybean meal, optimized mass transfer	separators or cross flow filtration	ion exchangers, crystallization or direct spray drying	>100 g/L L-lysine in 80 hrs

alternative process: cell reactor



D-α-amino-ε-caprolactam (ACL)

L-α-amino-ε-caprolactam

L-lysine

1 ACL racemase *Achromobacter obae*

2 ACL hydrolase *Cryptococcus laurentii*

Aspartame™, L-phenylalanine, and L-aspartic acid

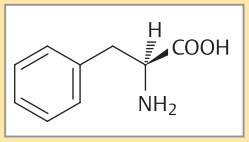
General. Aspartame™ (L-α-aspartyl-L-phenylalanine-methylester) is a synthetic low-calorie sweetener that is about 200 times sweeter than sucrose (→180). It was discovered in 1965 at G. D. Searle and cleared in 1981 by the FDA as a food additive. It is produced on a scale of ca. 20,000 $t\,y^{-1}$ (2009); the market leader is Ajinomoto. The starting materials for aspartame are L-aspartic acid and L-phenylalanine. Chemical synthesis requires the use of several protecting groups and is not competitive with enzymatic synthesis.

L-Aspartic acid can be isolated by extraction from protein hydrolysates. The preferred synthesis, however, is the addition of ammonia to fumaric acid by the enzyme aspartase in whole cells of *Escherichia coli*. Usually, a cell reactor (→102) is used, with the bacteria immobilized to κ-carrageenan or polyacrylamide. The productivity of this system is around 140 $g\,L^{-1}\,h^{-1}$ and the operational stability (half life of the catalyst) up to 2 y. Using lyophilized cells after induction, yields were up to 166 $g\,L^{-1}$. The formation of aspartase in *E. coli* K12 could be increased 30 fold when a plasmid containing the gene coding for aspartase (*aspA*) was expressed. Compared to a cell reactor, fermentation processes are not competitive, even when high-performance mutants are used.

L-Phenylalanine. In the past, industrial production was usually based on enzyme reactors using readily available chemical building blocks. Recently, fermentation processes based on high-performance mutants have become competitive. The availability and price of the synthetic precursors compared to the space-time yields of the fermentation process are the decisive factors in the economic preference. Using enzyme reactors, the best results were obtained by the addition of ammonia to *trans*-cinnamic acid, using phenylalanine ammonia lyase from *Rhodotorula glutinis*. In a cell reactor with immobilized microorganism, yields are ca. 50 $g\,L^{-1}$ at 83 % turnover. Cleavage of D,L-5-benzylhydantoin by L-hydantoinase and L-N-carbamoylase from *Flavobacterium ammoniagenes* was also promising (→132). For current fermentation processes, high-performance mutants of *E. coli* are used. The biosynthesis of L-phenylalanine proceeds from the precursors erythrose-4-phosphate and phosphoenolpyruvate via the intermediates shikimic acid, chorismic acid, and prephenic acid. This biosynthetic pathway includes branches leading to L-tyrosine and L-tryptophan, and many enzymes on this pathway are highly regulated. Consequently, auxotrophic mutants are used for the production of L-phenylalanine. As all production strains are auxotrophic for L-tyrosine, another advantage is that growth can be controlled by the addition of L-tyrosine to the medium. Most genes of the pathway have been cloned, and genetic methods have also been used to produce strains with derepressed production. Yields of 50 $g\,L^{-1}$ after 60 hrs have been reported using recombinant strains of *E. coli*. Fermentation is carried out in a fed-batch mode (→92), the cells are removed by filtration, and product is usually recovered by concentrating the broth by ultrafiltration, followed by adsorption chromatography and crystallization.

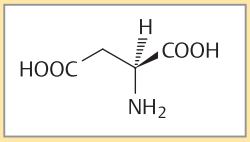
Aspartame™. The chemical synthesis of Aspartame™ from its two constituent amino acids requires the use of protecting groups and their later removal. Compared to this multi-step process, the currently used enzymatic production process is much simpler: only the amino group of L-aspartic acid must be protected, and an enzyme is used to catalyze amidation of the α-carboxy group of L-Z aspartic acid (the isomeric L-β-aspartyl-L-phenylalanyl-methylester tastes bitter) with L-phenylalanyl methylester. The regioselectivity of this enzyme even allows racemic phenylalanine methylesters to be used in this reaction. The preferred enzyme for this reaction is immobilized thermolysin from *Bacillus stearothermophilus*, and the solvent is i-amyl alcohol. Thermolysin is quite temperature stable and allows for a process temperature of 70 °C, leading to very high space-time yields of 30 $g\,L^{-1}\,h^{-1}$. The reaction product is of high purity and is further purified by ion-exchange chromatography. Aspartame is just one of a series of natural low-calorie sweeteners which were recently introduced. Other examples are sucralin (a trichloro-galactosucrose), stevioside (a glycosylated diterpene), thaumatin or monellin (glycosylated proteins or peptides, respectively), and Advantame™, a modified aspartyl-phenyl-alanine dipeptide developed by Ajinomoto.

L-Phenylalanine



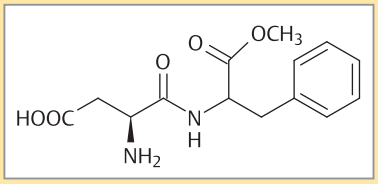
$C_9H_{11}NO_2$	
M_R	165.19
decomp.	310 – 312 °C
CAS	63-91-2

L-Aspartic acid



$C_4H_7NO_4$	
M_R	133.10
CAS	56-84-8

α-Aspartame



$C_{14}H_{18}N_2O_5$		* accep-
M_R	294.31	table
ADI*	40 mg/kg	daily
CAS	22389-47-0	intake

Production of L-phenylalanine

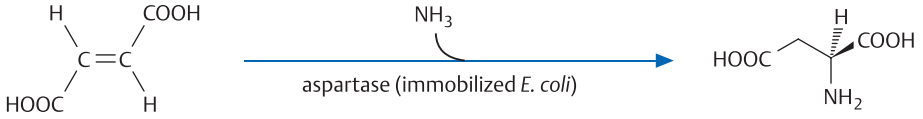
preferred process: fermentation		
high-performance mutants of *E. coli*	C-source: glucose Recovery: ultrafiltration, ion exchange chromatography, crystallization	50 g L^{-1} in 60 h
alternatively: enzyme reactor		
addition of ammonia to *trans*-cinnamic acid	carrier-bound L-phenylalanine-ammonia lyase from *Rhodotorula glutinis*	~ 50 g L^{-1}

Production of L-aspartic acid

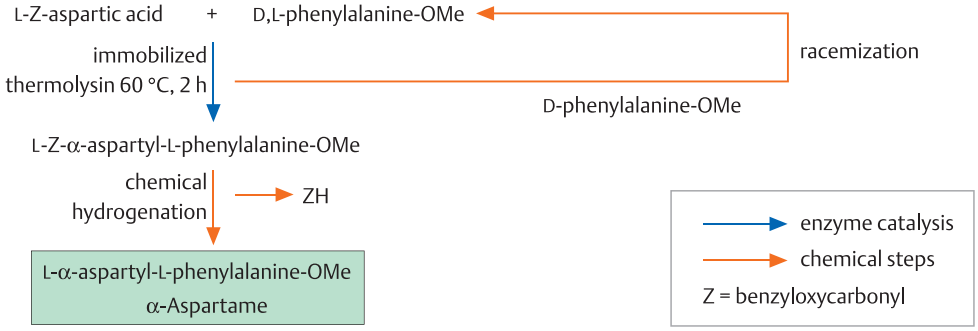
multi-stage cell reactor
immobilized *E. coli* cells; 1.2 M ammonium fumarate pH 8.5, 37 °C

recovery
with H_2SO_4 at 15 °C to pH 2.8: aspartic acid precipitates

90 – 95 % yield
half-life of *E. coli* cells: 2 yr



Production of α-Aspartame



sweetener	chemical constitution	relative sweetness
saccharose	disaccharide	1
cyclamate	synthetic cyclohexyl sulfamide, Na salt	40
α-Aspartame	dipeptide methylester	200
stevioside	glycosylated diterpene	300
saccharin	synthetic 2-sulfobenzoic acid imide, Na salt	450
sucralose	chlorinated disaccharide	600
thaumatin	nonglycosylated protein from 208 amino acids	2 500
monellin	nonglycosylated protein, two peptide chains of 44 and 50 amino acids	2 500
Advantame	Aspartame derivative	20 000

Amino acids *via* enzymatic transformation

General. As was shown before for several examples (lysine (→128), aspartic acid (→130), phenylalanine (→130)), chiral amino acids can be prepared by the enzymatic transformation of racemic building blocks. The advantage of enzyme processes over fermentation procedures rests in the possibility of preparing non-proteinogenic and even non-natural amino acids. In most enzymatic transformations of this kind, hydrolases and racemic precursors are used. Examples are esterases, aminoacylases, amidases, and hydantoinases/carbamoylases. The disadvantage of this process lies in the need to reracemize the "wrong" enantiomer in a coupled reaction, in order to feed it again into the reaction. Consequently, addition reactions based on lyases and redox reactions based on oxidoreductases have been intensively studied, since they lead to only one enantiomer.

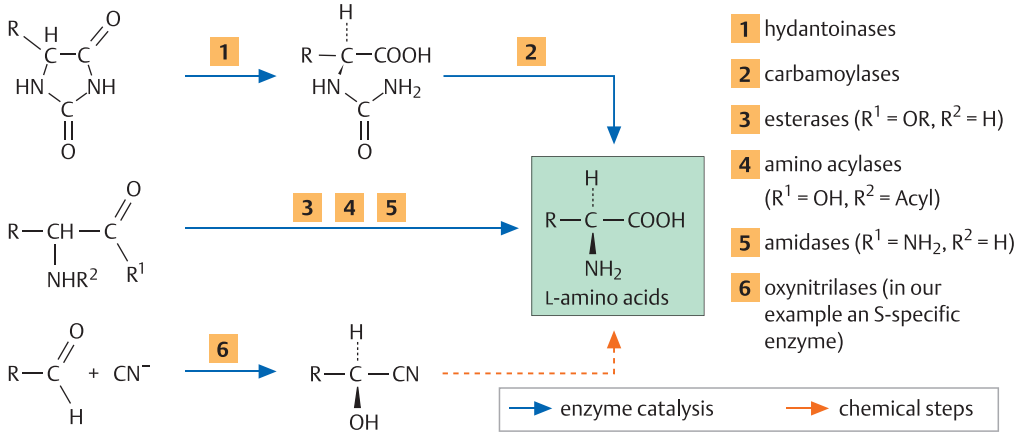
Enantioselective hydrolysis. To date, enzyme reactors (→102, 164) based on aminoacylases and hydantoinases are the most advanced and some have already been industrialized. In the aminoacylase reaction, carrier-bound enzymes (e. g., from *Aspergillus oryzae* or *Bacillus thermoglucosidius*) hydrolyse racemic N-acyl amino acids. Only the L-enantiomer is hydrolyzed. The N-acyl-D-amino acid remains in the reaction mixture, from which the L-amino acid is separated by crystallization. After the wrong enantiomer has been reracemized, often in a thermic reaction, it is combined with new racemic precursor and fed into the reactor again. Using this technology, several 100 t y^{-1} of L-methionine, L-tyrosine, L-proline, and L-valine are produced for clinical use (mainly infusions). Although D-amino acids can also be produced by this procedure, hydantoins, which can be easily prepared chemically, and hydantoinases, which can be isolated in many variations from microorganisms, offer the better choice for the preparation of non-proteinogenic and non-natural amino acids. In this process, racemic hydantoins are cleaved with hydantoinases of the desired specificity. N-carbamoylated amino acids are formed as the primary product and can be hydrolysed to the desired chiral amino acid. The "wrong" hydantoin is then racemized at pH 8.5 and becomes available for a new hydrolysis cycle. This reaction sequence, when integrated in a suitable host organism such as *Escherichia coli*, is used, e. g., for the industrial manufacture of R-phenylglycine and R-4-hydroxyphenyl glycine, the side chains of Ampicillin and Amoxicillin (semisynthetic penicillins).

Enantioselective addition reactions. Oxynitrilases occur mainly in plants. They exhibit R or S selectivity and so do not lead to formation of a byproduct of undesired chirality. Both types of oxynitrilases have been cloned and expressed in *E. coli* or yeasts and thus are readily available. Examples are R-oxynitrilase from Manihot and the S-oxynitrilase from almonds. The crystal structure of both enzymes has been solved, and researchers are now using protein engineering to enhance their substrate specificity for their eventual use in various industrial applications.

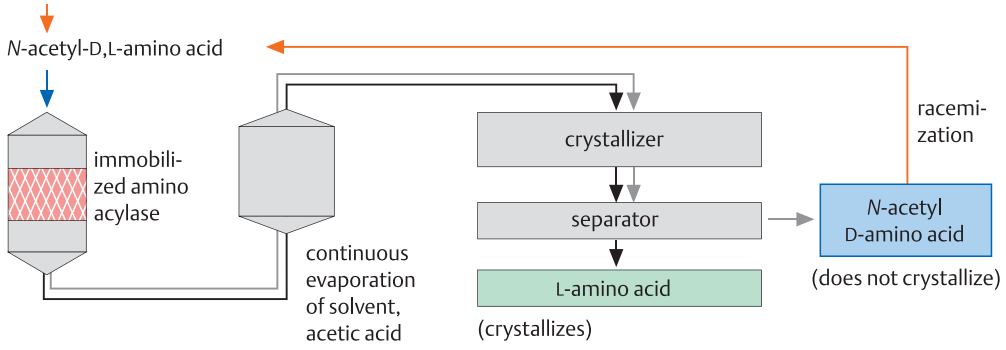
Enantioselective redox reactions. The example of the stereoselective synthesis of L-leucine from synthetic α-oxo caproic acid makes it evident that reductive amination of the oxo compound by L-leucine dehydrogenase from *Bacillus sp.* requires not only NH_3, but also NADH. Since NADH is expensive, its regeneration is imperative for achieving an economic process. An elegant older procedure is based on the use of formate dehydrogenase from *Candida boidinii*, since the reaction product CO_2 evaporates from the mixture and thus shifts the reaction equilibrium towards L-leucine. If polyethylene glycol (PEG) is bound to NADH, the cofactor remains functional and is retained in an enzyme membrane reactor. With this technology, up to 6×10^5 mole equivalents of product were obtained per mole of NADH-PEG consumed. More recently, NADH or NADPH regeneration is achieved by auxiliary enzymes such as alcohol or glucose dehydrogenases (→170), and the process is done in batch mode with the substrate-converting redox enzyme. The drawback of this procedure is that the by-products such as acetaldehyde or gluconic acid must be removed on product recovery. Most enzymes used in such reactions are cloned and often modified via protein engineering (→198) to increase their stability and specificity (e. g., for the regeneration of NADPH instead of NADH). Often, the whole reaction sequence is integrated in a host organism such as *E. coli*.

Major reaction types

reaction type	enzyme type	comments
hydrolysis of racemic precursors	hydrolases	preferred, simple reaction, but expensive, since the "wrong" enantiomer must be reracemized
addition of HCN or ammonia to carbonyl compounds	lyases	simple, quantitative reaction, but limited choice of enzymes
reductive amination of α-oxo carboxylic acid	dehydrogenases	only one enantiomer is formed, but auxiliary enzymes and cofactors render the process expensive



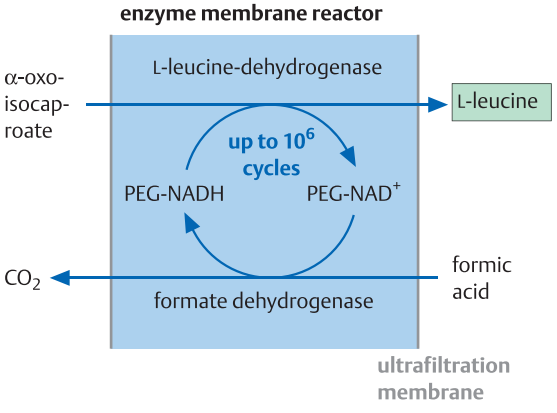
Enantioselective hydrolysis of *N*-acyl-D,L-amino acids



Reductive amination of α-oxocarbonic acids

the broad substrate specificity of leucine dehydrogenase allows the reductive amination of pseudosubstrates to non-proteinogenic amino acids, e.g., t-leucine

leucine dehydrogenase from *Bacillus sphaericus*	relative activity	K_m
α-oxoisocaproate	100	0.31
α-oxoisovaleriate	126	1.4
α-oxovaleriate	76	1.7
α-oxobutyrate	57	7.7
α-oxocaproate	46	7.0



PEG-NADH is a synthetic derivative of NADH (M_R ~ 3 000). It is accepted by many dehydrogenases as a cofactor and retained by ultrafiltration membranes

Vitamins

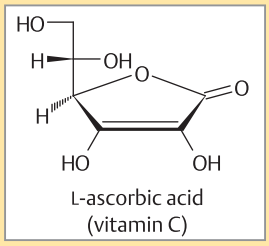
General. Vitamins are used as additives in medical preparations and animal feed. Most vitamins are manufactured by chemical synthesis or by extraction of plant material. Biotechnological processes are used in the syntheses of vitamins B_2, B_{12}, and C. The production volume of vitamins produced by fermentation exceeds 200,000 t/y (world), with vitamin C taking the largest share.

Vitamin B2 (riboflavin). The riboflavin derivatives FAD and FMN are cofactors of many redox enzymes. Free riboflavin occurs only in milk. Animals fed a riboflavin-deficient diet develop dermatitis, growth deficiencies, and damage to the eyes. Biosynthesis proceeds on a complex pathway starting from guanosine triphosphate and ribulose-5-phosphate. Riboflavin can be produced by chemical synthesis, but recently two fermentation based pathways have shown the best cost efficiency, and production is 10,000 t/y. In a process using genetically engineered mutants of the filamentous fungus *Ashbya gossypii* in fed-batch fermentations (carbon source: plant oil, nitrogen source: soymeal) yields reach > 20 g L^{-1} in 72 h. Recovery starts with autolysis of cells after heat shock treatment and includes cell separation and chromatographic procedures. In a second process with *Bacillus subtilis*, mutants are used which were generated by random mutagenesis and genetic engineering. In developing such mutants, care was taken to overproduce both building blocks, GTP and ribulose-5-P, simultaneously.

Vitamin B12 (cyanocobalamin). Vitamin B_{12} is required as an enzyme cofactor in several methylation and isomerization reactions. Vitamin B_{12} deficiencies lead to anemic conditions, in particular to pernicious anemia. Although it is used in medical preparations (mainly for liver protection) and in diets, half the world production of some 20 t is added to animal feed. Biosynthesis starts from succinoyl CoA and glycine, which condense to 5-amino 4-oxo valeric acid (δ-amino levulinic acid). In an additional 30 steps, which may vary among different organisms, 5'-deoxyadenosyl cobalamin is formed. For the manufacture of vitamin B_{12}, a fermentation process based on *Pseudomonas denitrificans* is dominant. Using a medium with molasses (which contains glycinbetain, an important cosubstrate) as a carbon source, ammonia, and cobalt salts, and using 5,6 dimethylbenzimidazol as a precursor compound, yields may reach 200 mg L^{-1} after 120 h. In *Propionibacterium shermanii*, a human skin-associated microorganism used in probiotics, all 22 genes of the vitamin B_{12} biosynthetic pathway were cloned, and modern production strains have already been optimized through metabolic engineering (→318).

Vitamin C (L-ascorbic acid). Ascorbic acid is often called a "physiological reducing agent" and participates in the reductive elimination of active oxygen species, e. g., oxygen radicals. In addition, it participates as a cofactor in several enzyme reactions. Vitamin C deficiency results in damage to the skin and the blood vessels (scurvy). It is sold as a vitamin preparation and is also added to numerous foods and drinks both as a vitamin and as an antioxidant. Annual production exceeds 100,000 t (world) and is mostly done in China. The manufacturing process is based on a chemo-enzymatic synthesis starting from D-glucose, which is chemically hydrogenated to D-sorbit. The traditional route is the Reichstein-Gruessner synthesis or modifications thereof, which comprises subterminal oxidation of D-sorbitol to L-sorbose with *Gluconobacter oxydans* and chemical oxidation to 2-keto-L-gulonic acid (2-KGA), requiring protection/deprotection of the chiral hydroxy groups. The overall yield of this process is ca. 66 %. Oxidation is done continuously or batchwise, requiring immobilized cells and strong aeration. Yields are quantitative after 24 h. In the 2-KGA process, L-sorbose is oxidized by *Ketogulonicigenium vulgare* directly to 2-KGA in high space-time yields (94 % yield after 26 h), which lactonizes easily to L-ascorbic acid. For unknown reasons, a second microorganism (such as *Bacillus megaterium*) is required for this process, but *Gluconobacter oxydans* required for L-sorbose oxidation can also be used, reducing the overall steps to three. Cloning and expressing all necessary enzymes in a suitable host organism is the logical next step, first attempted as early as 1993 by Genentech using *Erwinia herbicola,* but the glucose tolerance of the recombinant production strain was too low, resulting in unsatisfactory space-time yields.

Vitamins

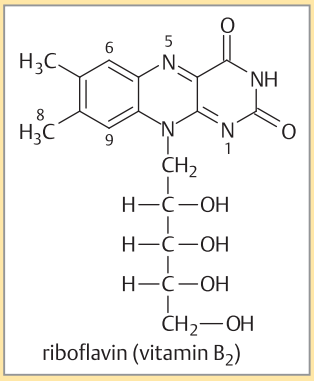


L-ascorbic acid
(vitamin C)

$C_6H_8O_6$

M_R 176.13

CAS 50-81-7



riboflavin (vitamin B_2)

$C_{17}H_{20}N_4O_6$

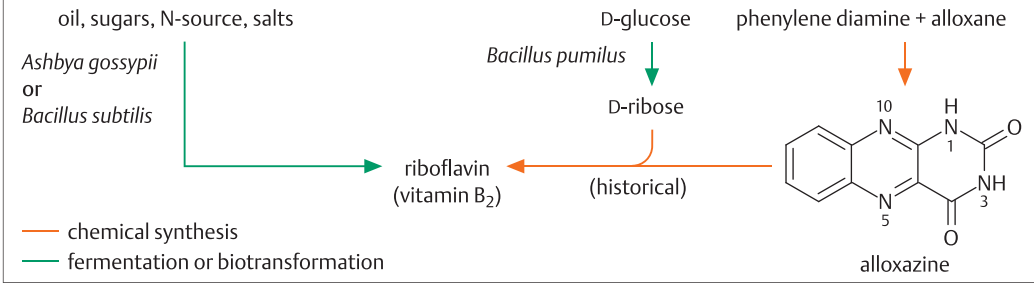
M_R 376.36

CAS 83-88-5

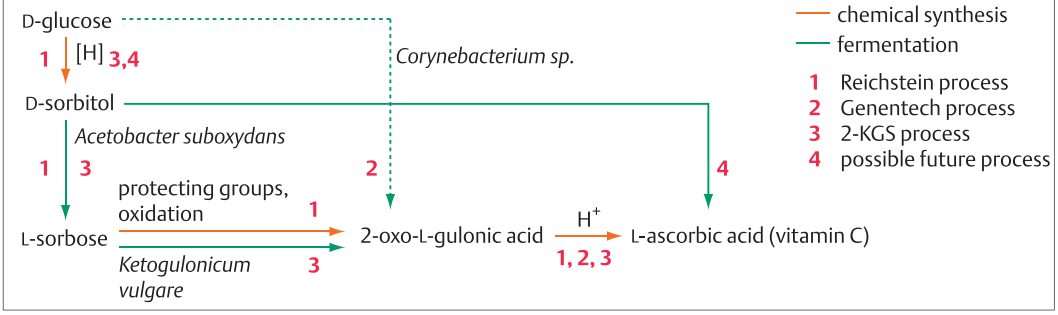
Manufacturing processes and market volume

vitamin		quantity (t), ~ 2010	manufacturing process	applications
A	β-carotene	3 000	chemical synthesis	animal nutrition, colorant
B_1	thiamine	3 000	chemical synthesis	health
B_2	riboflavin	6 000	fermentation (30 %), chemoenzymatic synthesis (50 %), chemical synthesis (20 %)	health, animal nutrition
B_6	pyridoxine	3 000	chemical synthesis	health, animal nutrition
B_{12}	cyanocobalamine	20	fermentation	health, animal nutrition
C	ascorbic acid	>100 000	chemo-fermentative synthesis	health, food additive, animal nutrition
D_2	calciferol	40 000	photochemical synthesis from ergosterol	health
E	α-tocopherol		extraction from plant oil, algae	health

Biotechnological routes to riboflavin



Synthesis of L-ascorbic acid



Nucleosides and nucleotides

General. About 50 years ago, 5'-nucleotides were discovered in Japan as taste-enhancing components in dried mushrooms and dried fish. If even very low quantities (0.0005–0.001 %) are added to soups and sauces, their taste is significantly enhanced, and undesired off-flavors, e. g., the metallic taste of canned foods, are suppressed. Addition of Na-L-glutamate (→126) enhances this effect. Inosine-5'-monophosphate (IMP), guanosine-5'-monophosphate (GMP) and xanthosine-5'-monophosphate (XMP) exhibit the strongest activity, and the 2'- and 3'-isomers, pyrimidine nucleotides, and nucleosides have no effect. Adenosine-5-hibit the strongedeoxy IMP, and deoxy GMP are less active than 5'-IMP or 5'-GMP. Since 1961, 5'-IMP and 5'-GMP have been produced on an industrial scale, at present (2014) > 10,000 t y^{-1}. The producers are predominantly located in Asia.

Manufacturing process. Four processes are used: 1) enzymatic hydrolysis of RNA, 2) combination of fermentative production of inosine or guanosine, and their chemical phosphorylation, 3) direct production of 5'-IMP by fermentation, and 4) direct production of 5'-XMP by fermentation, and its enzymatic conversion to 5'-GMP.

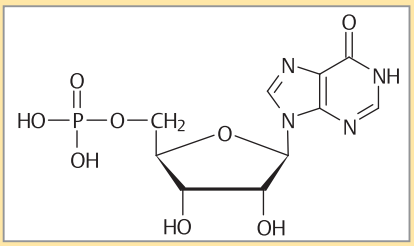
Enzymatic hydrolysis of RNA. Yeasts show the most favorable RNA/DNA ratio and thus are preferentially used for the production of RNA. If *Candida utilis* (→14) is grown on molasses or pulp, using media with a low C/N-ratio, cells contain 10–15 % RNA (dry weight) in the early exponential phase. This amount can be further increased by adding Zn^{2+} and phosphate. The aerobic fermentation is usually carried out continuously on a large scale, using airlift bioreactors (→96). After removal of the cell mass, the RNA is extracted with a hot, alkaline NaCl solution (5–20 % NaCl, 100 °C, 8 h), and precipitated with HCl or ethanol. For enzymatic hydrolysis of the RNA, nuclease P_1 preparations from *Penicillium citrinum* are used which do not contain unspecific 5canucleotidases or phosphatases. The final isolation of 5'-IMP and 5'-GMP proceeds by adsorption onto activated charcoal, ion-exchange chromatography, and crystallization.

Fermentation of 5llizati The classical protocol is based on the formation of inosine by fermentation with *Bacillus* sp. and other Gram-positive microorganisms. Inosine is secreted into the medium and can be precipitated at pH 11. Fermentation of adenine-auxotrophic mutants whose transport properties have been improved by genetic engineering may result in yields of 35 g L^{-1}. Recrystallized inosine is then transformed to 5'-IMP with PCl_3 in suitable solvents. Currently preferred procedures lead directly to 5'-IMP. *Corynebacterium ammoniagenes* block mutants are no longer repressed by nucleosides, are not sensitive to the Mn^{2+} content of the medium, and secrete 5'-IMP without hydrolizing it. In a recent procedure, permeabilized *C. ammoniagenes* cells are co-incubated with *E. coli* mutants overexpressing an inosine kinase gene. The resulting inosine kinase forms 5'-IMP from inosine and ATP, with *C. ammoniagenes* providing the inosine and regenerating ATP from ADP and inexpensive inorganic phosphate. Another Japanese process uses recombinant *E. coli* cells forming a phosphatase/phospho-transferase from *Escherichia blattae*. Using protein engineering (→198), the phosphatase activity of this enzyme was largely removed, and the remaining phosphotransferase activity phosphorylates inosine with inorganic phosphate in very high yields.

Production of 5 in ver The preferred processes are 1) the production of guanosine by fermentation, followed by chemical phosphorylation, and 2) production of XMP by fermentation and its enzymatic conversion to GMP. GMP formation via 5-aminoimidazole carboxamide-1-riboside (AICA-riboside), obtained by purine auxotrophic strains of *Bacillus megaterium*, and followed by chemical transformation into 5'-GMP, has also been reported, but is not industrially used at present.

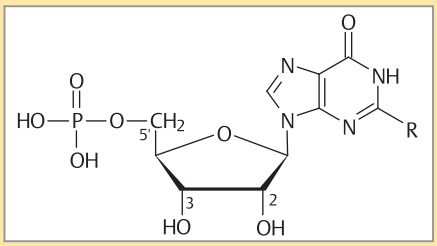
Other nucleotides. Nucleotides such as ATP, cAMP, NAD^+/NADH, $NADP^+$/NADPH, FAD, coenzyme A, and nucleotide sugars are important biochemicals, which have also been used in special biotransformation processes. They are prepared either by fermentation using block mutants or by enzymatic processes starting from chemical precursors. As an example, both NAD^+ and coenzyme A can be prepared by fermentation using block mutants of *Brevibacterium* (*Corynebacterium*) *ammoniagenes* in yields of ca. 2 g L^{-1}.

Inosine 5'-monophosphate



$C_{10}H_{13}N_4O_8P$
M_R 348.21
CAS 131-99-7

Guanosine 5'-monophosphate



R = NH_2: guanosine 5'-monophosphate (5'-GMP)
$C_{10}H_{14}N_5O_8P$ M_R 363.22 CAS 85-32-5
R = OH : xanthosine 5'-monophosphate (5'-XMP)

Manufacture and market volume

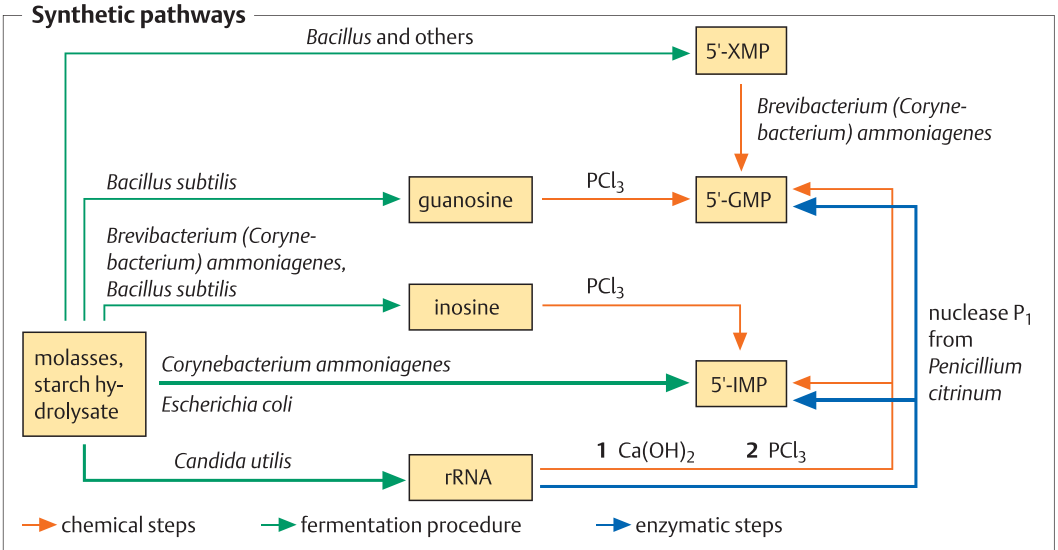
nucleoside	market volume (t)	manufacturing process	applications
5'-IMP, 5'-GMP	>10000	enzymatic hydrolysis of yeast RNA fermentation of inosine/guanosine and chemical phosphorylation or direct fermentation of 5'-IMP	flavor enhancer flavor enhancer
inosine	25	fermentation	medical therapy
orotic acid	20	fermentation	liver diseases
adenine, adenosine, ATP	20	fermentation, enzymatic synthesis	medical therapy

DNA-/RNA-content of microorganisms

	bacteria	yeast	fungi
DNA*	0.37 – 4.5	0.03 – 0.52	0.15 – 3.3
RNA*	5 – 25	2.5 – 15**	0.7 – 28

* % of dry cell mass
** of which 75 – 80 % is rRNA

Synthetic pathways



fermentation	recovery		chemical synthesis
high-performace strains of *B.* (*C.*) *ammoniagenes*; starch hydrolysate, raw RNA, NH_2; pH 6, 30 °C	precipitate at pH 11; recrystallize	> 30 g L^{-1} inosine after 42 h	of 5'-IMP: PCl_3 in trialkylphosphate, yield > 90 %

Bio-Ethanol

General. Ethanol is an important industrial solvent and a starting material for the synthesis of organic chemicals such as ethylene and polyethylene. In many countries, ethanol is added as a biofuel to gasoline (E10 and E15: 10 or 15 % by volume ethanol). "Flex-fuel" vehicles tolerate higher ethanol concentrations. In 2012, about 85 billion L of bioethanol (~67 million tons) were produced, 77 % in the Americas. 95 % of the global production is based on biotechnology using glucose or saccharose as a carbon source, and yeast or bacteria as a biocatalyst. As of 2014, the production of bioethanol is economically competitive to petrochemistry only if very low prices for the C-source (→328) are obtainable or if oil prices rise or are regulated by the government. In Brazil, the C-source is saccharose (cane molasses), and in the USA it is glucose obtained from corn starch (→178). Since corn is also used as an animal feed and for human nutrition, a potential conflict about the use of this raw material ("eat or drive") is countered by intense efforts to produce bioethanol from biomass (straw, switch-grass, etc.) (→182) in biorefineries (→330).

Organisms and biosynthesis. The most important organism used in producing ethanol is bakers' yeast (*Saccharomyces cerevisiae*) (→14). By glycolysis, it forms two moles of ethanol per mole of glucose. *Zymomonas mobilis*, a bacterium isolated from agaves, achieves the same molar yield, but the synthesis is based on the ketodeoxyphosphogluconate (KDPG) metabolic pathway. Both organisms can metabolize saccharose but not starch, as they do not possess starch-depolymerizing enzymes. Thus, if starch is to be used as a C-source, depolymerization to glucose must be done prior to fermentation.

Fermentation and recovery. The production of bioethanol is usually carried out in the fed-batch mode using large-scale bioreactors (→96) of up to 500m^3 capacity. The organism of choice is *Saccharomyces cerevisiae*, which is less sensitive to *Lactobacillus* infections than *Zymomonas mobilis,* even under non-sterile conditions. For cultivation, a sugar medium enriched with N-sources and minerals is used. After an aerobic growth phase, aeration is stopped, and after around 20 h ethanol production reaches 90 % of the theoretical maximum. Glucose concentrations >0.1 % lead to inhibition of the process by catabolite repression; this inhibition is prevented by a continuous or semi-continuous feeding of sugar (fed-batch) (→92). Since ethanol concentrations >8 % (which are usually reached after 72h) inhibit yeast metabolism, the ethanol-containing broth is removed. 95 % ethanol, which can be used in cars, is usually isolated by azeotropic distillation. Absolute ethanol can be prepared by extractive distillation, molecular sieves, or membrane technology (pervaporation). In several processes, e. g., the Melle-Boinot process, the cells obtained by a separator during broth removal are recycled and serve as a pre-culture for the next batch; this procedure reduces the overall fermentation time. Continuous ethanol fermentation methods have been developed but are technically more demanding. A cell reactor (→102) with immobilized yeast has been explored for production but are not currently used.

Economic considerations. At present, > 700 fermentation plants produce bioethanol. In Brazil, a "Proalcool" program was initiated in 1975, aimed at reducing oil imports. In over 100 fermentation plants, around 21 billion L/y (2012) of ethanol are produced from sugarcane molasses using simple technology such as batch fermentations or distillation). In the USA, automotive gasoline enriched bio-ethanol ("gasohol") has been produced since 1975 (2012: ca. 50 billion L/y) from cornstarch. The technology is more demanding (e. g., fed batch, pervaporation). Fermentation solids (cell mass and solid media components) are sold as animal feeds.

Ethanol from biomass (→128, 328). The first step in this process is a physical (e. g., steam) or chemical (e. g., phosphoric acid, ammonia) disintegration, followed by enzymatic hydrolysis using cellulases and hemicellulases (→182). In this process, enzyme inhibitors such as furfural are formed, and hemicelluloses are hydrolyzed to pentoses such as xylose and arabinose that cannot be metabolized by baker's yeast. As a consequence, emphasis is on efficient and inexpensive pulping procedures, on cheap and inhibitor-resistant hydrolytic enzymes, e. g., from *Trichoderma reese*, and robust, recombinant yeast or bacterial strains (*S. cerevisiae, Kluyveromyces marxianus, Pichia stipitis, Z. mobilis, E. coli*) which can metabolize both glucose and pentoses.

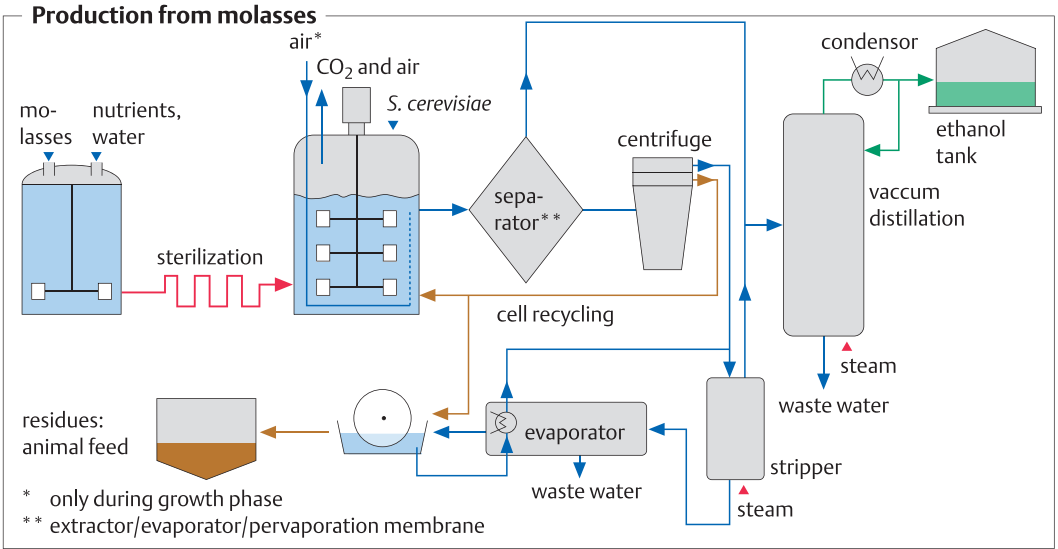
Ethanol

C_2H_6O

M_R	46.07	b. p.	78.32 °C	Industrial production:
D.	0.79367 (15 °C)	CAS	64-17-5	mainly addition of water to ethylene using catalysts

$H_3C—CH_2—OH$

Production from molasses



* only during growth phase
** extractor/evaporator/pervaporation membrane

Fermentation and recovery

nutrients
sugarcane juice or beet molasses or saccharified starch, minerals

cell reactor
immobilized cells, continuous process

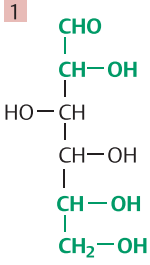
bioreactor
up to 500 m^3 batch or fed-batch process

ethanol separation
distillation or pervaporation

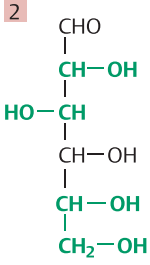
waste material
animal feed

Ethanol

1



2



origin of the C atoms of ethanol during metabolism of glucose with

1 *Saccharomyces cerevisiae*	glycolysis
2 *Zymomonas mobilis*	KDPG pathway

Conversion of biomass or corn starch to bioethanol

a lignocellulose hydrolysis with phosphoric acid

Lignocellulose → dilute acid hydrolysis → liquefaction (+ cellulases and glucosidases) → fermentation* (+ cellulases and glucosidases) → distillation

stainless steel reactor

b corn starch to ethanol

Corn → steam cooker → liquefaction (+ amylase an glucoamylase) → fermentation (+ amylase an glucoamylase) → distillation

stainless steel reactor

*co-fermentation of glucose (from cellulose) and pentoses (from hemicellulose)

1-Butanol

General. 1-Butanol (2011: world production ca. 3 million t) is an important solvent for automobile paints, a base chemical for ester formation (e. g., for butyl cellulose) and a biofuel. Acetone, the side product of 1-butanol fermentation (2009 world production: ca. 6.7 million t) is also used as a solvent. During World War I, it came in great demand for the production of cordite, an explosive used by the British navy. Both compounds are presently produced from petrochemical raw materials, but before ca. 1950 they were mostly obtained by fermentation using *Clostridium* bacteria and starch or molasses as a carbon source, an industrial process pioneered in 1915 by the Russian/British chemist Chaim Weizmann (who later became the first president of Israel). Due to progress in molecular genetics and process technology, the production of either solvent by fermentation might eventually become economically attractive again and is being investigated as a reserve technology.

Organisms and biosynthesis. Among the few anaerobic bacteria that can form acetone and 1-butanol, the genus *Clostridium* is the most important. During fermentation, a shift from the formation of butyric and acetic acid ("acetogenesis") to the formation of butanol occurs at the end of cell growth and is accompanied by a decrease in pH to values < 5.0. The composition of the product mixture varies from species to species. The best-studied organism is *Clostridium acetobutylicum*, which also shows the highest tolerance to the cell-toxic solvents formed. It can form up to 38 g of 1-butanol and acetone from 100 g glucose in a ratio of 3:1. A side product is ethanol ("ABE fermentation"). Many Clostridia synthesize amylases, amyloglucosidases, and other extracellular hydrolases and thus are capable of metabolizing inexpensive carbon sources (→328) such as starch. The use of glucose and pentoses derived from biomass is also studied, as is the use of lactose (whey). The enzymes participating in the biosynthesis of both solvents have been well studied and their genes have been cloned. Pyruvate is formed from glucose by glycolysis. In the presence of pyruvate/ferredoxin oxidoreductase, pyruvate undergoes oxidative decarboxylation to acetyl CoA, which is further reduced to several C_2-, C_3-, or C_4-metabolites using mainly NADH from glycolysis for reduction. A hydrogenase that is also present transfers some of the electrons to protons with the formation of hydrogen. The regulation of these enzymes is being intensively studied with the goal of influencing the yield and the composition of solvents, as is optimization of the pathway through metabolic engineering. The genome of *C. acetobutylicum* has been completely sequenced, and the organism is promising for genetic engineering, since shuttle vectors for *Escherichia coli* and *Bacillus subtilis* and are available specific phages and transposons have been identified. The yield of butanol, at the expense of acetone, could already be enhanced to 15 % (w/v) by engineering production strains which turn acetone into the "butter flavor" acetoine.

Fermentation and recovery. For more than 40 years, the production of acetone and 1-butanol has been carried out on an industrial scale using *C. acetobutylicum* and batch fermenters of > 100 m^3 volume. Substrate costs in this process were ca. 60 %, energy costs for product distillation ca. 12 %. Decisive parameters for renewed use of the fermentation route are the yield of product from raw material (kg solvent kg^{-1} sugar) and the productivity of the process (g solvent L^{-1} h^{-1}). Modern process developments thus aim at two-stage processes with cell recycling, biofilm processes with integrated macroporous resins and improved recovery of the solvents by pervaporation.

Economic considerations. At present, the batch process based on cornstarch or molasses as the carbon source, which has been used in the USA and South Africa for over 40 years, is not competitive with petrochemical-based synthetic routes. Only in China have larger plants been built, with up to 200 m^3 fermentation volume that produce some 30,000 t of 1-Butanol. The C-source is corn starch, and strain *Clostridium acetobutylicum* EA2018 (improved by metabolic engineering (→318), in particular by deletion of gene *adc* coding for acetoacetyldecarboxylase) is used. Yields are at 14 g L^{-1}, and product composition is high in 1-butanol with 2:7:1 (ABE). Smaller pilot plants in the USA (Cobalt Technologies) use continuously cultured *C. acetobutylicum* with glucose from hydrolyzed biomass. High yields were also reported from strains of *E. coli* modified by metabolic engineering (→318).

1-Butanol

$C_4H_{10}O$ — $H_3C-CH_2-CH_2-CH_2-OH$

m.w.	74.12
d	0.813
m. p.	– 90 °C
b. p.	117 – 118 °C
CAS	71-36-3

chemical synthesis of 1-butanol:
hydroformylation of propene and hydrogenation

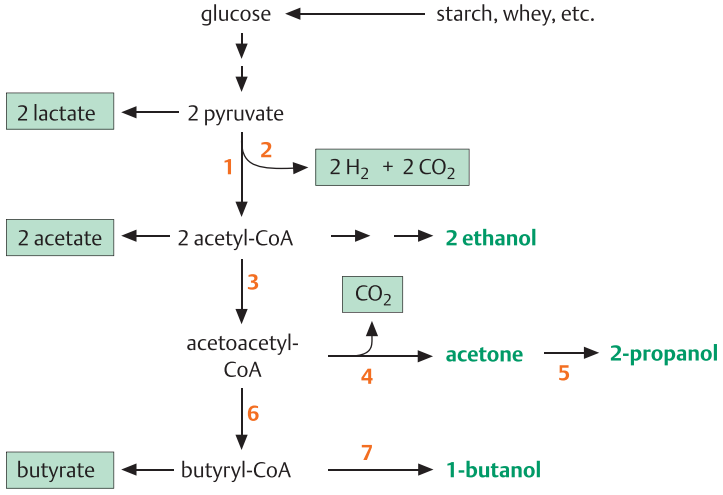
Acetone

C_3H_6O — $H_3C-CO-CH_3$

M_R	58.08
d	0.7908
m. p.	– 95 °C
b. p.	56 °C
CAS	67-64-1

chemical synthesis of acetone:
catalytic dehydrogenation of 2-propanol, direct oxidation of propene, cleavage of cumol hydroperoxide

Biosynthesis



1 pyruvate: ferredoxin oxidoreductase
2 hydrogenase
3 thiolase
4 CoA-transferase/ acetoacetate decarboxylase
5 aldehyde-/alcohol dehydrogenase
6 two dehydrogenases, crotonase
7 butyraldehyde- and butanol dehydrogenase

the genes coding for all enzymes 1 – 7 have been cloned

Fermentation and recovery

spore suspension
suspension in soil

shake flask culture
anaerobic, 24 hrs at 37 °C

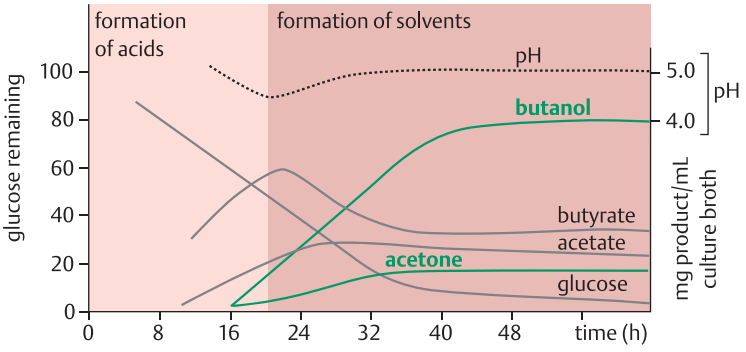
15 m³ bioreactor
under CO_2 layer
18 hrs at 37 °C

700 m³ bioreactor
6 % molasses
50 – 60 hrs at 37 °C

continuous distillation
followed by fractionated distillation

~ 38 kg 1-butanol/ acetone (3:1)
100 kg^{-1} glucose

typical fermentation curve for *Clostridium acetobutylicum*



Yields of 1-butanol (2012, selection)

Microorganism	1-butanol (g/L)	ABE* (g/L)	scale (tons)	C-source	country
C. acetobutylicum EA2018	14	21	200000	corn starch	China
C. acetobutylicum BEKW	17	20	15	corn starch	Korea
C. acetobutylicum wild-type	no data		9000	lignocellulose	USA
E. coli BW	15	12**	15	corn starch	USA

*acetone, 1-butanol and ethanol **1-butanol and ethanol

Higher alcohols and alkenes

General. With rising oil prices, it has become more attractive to consider production of alcohols other than ethanol (→138) or 1-butanol (→140) from sugars through fermentation. Synthetic pathways may include both chemical and biological steps. In the latter case, metabolic engineering (→320) of production strains for the re-direction of metabolic flux to the desired end product is a crucial prerequisite. Higher alcohols are often members of a chemical product family and can also be dehydrated to alkenes, which, after polymerization, lead to biopolymers such as "bio-polyethylene." Several processes for the production of higher alcohols and alkenes from biomass sugars are presently under industrial development.

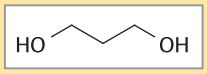
1,3-Propanediol (1,3-PD) is a building block for the synthesis of thermoplastic polyesters (→154). In 2007, DuPont introduced Sorona®, a polyester for the production of textiles and carpets made from terephthalic acid and 1,3 propanediol. The industrial production of 1,3-PD was a milestone in biotechnology, since for the first time a recombinant strain of *E. coli* with an artificial pathway composed of genes from *E. coli, S. cerevisiae and Klebsiella pneumoniae* was designed to be robust enough for industrial manufacture. Yields of 1,3-PD exceed 150 g/L. Decisive measures for the construction of this production strain were: 1) change of the metabolic flux at the triosephosphate isomerase branch of glycolysis by deletion of glycerokinase (*glpK*) and glycerol dehydrogenase (*gldA*), 2) Elimination of the phosphotransferase (PTS)-dependant glucose transport system, and 3) down-regulation of glyceraldehyde-3-phosphate dehydrogenase (*gap*).

2-Methyl-1-Propanol (Isobutanol) is a solvent that is synthesized by hydrocarbonylation of propene. It is used as a starting material for chemicals used in the lacquer and dye industry, but also as a "drop-in" chemical in fuels. A new procedure for production by fermentation is based on the use of metabolically engineered yeasts (Gevo). Isobutanol is also a starting material for the production of p-xylene, a "green" substitute for the synthesis of terephthalic acid (a component of PET polyesters) (→154). "Bio-terephthalic acid" competes with 2,5-furandicarboxylic acid (FDC), which can be chemically produced from D-glucose and is a potential dicarboxylic acid building block for the manufacture of PEF-type polyesters (Polyethylene Furane Polyesters).

2,3-Butane diol (2,3-BDO) is a chiral starting material for chemical synthesis and an interesting building block for the production of *cis*-1,3-butadiene, the key monomer of synthetic rubber. Butadiene can also be produced from biomass through dimerization of bioethanol at high temperature ("Lebedev-process") or by oxidative dehydrogenation of but-1-ene or but-2-ene (from 1-butanol or isobutanol). Another option is controlled dehydration of 2,3-BDO, a metabolic end product of the mixed acid fermentation pathway used by many anaerobic and facultative anaerobic microorganisms. For the industrial production of 2,3-BDO, engineered strains of *Enterobacter* or *Klebsiella* are particularly appropriate (Lanza). With suitable media and strictly controlled oxygen supply, yields exceeding 150 g/L in 38 hrs could be reached. The use of xylose as fermentation raw material (hemicelluloses from biomass) is also strongly promoted. Not yet completely solved is the reduction of by-product formation (ethanol, lactic acid), which at present must be separated from the product.

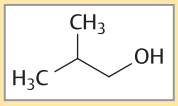
Isoprene, the monomeric building block of rubber, is a volatile chemical (b. p. 34 °C). It is formed in plants mostly through mevalonic acid, and in some plants and bacteria through a non-mevalonic pathway. Natural rubber synthesized by the rubber tree (*Hevea brasliensis*) consists mainly of *cis*-1,4-polyisoprene. A regio- and stereoselective chemical synthesis of this material requires special catalysts. Genencor und Dupont have expressed a plant isoprene synthase in *E. coli* and engineered the strain to overproduce the precursor 3,3-dimethylallyl pyrophosphate DMAPP, arriving at an industrial process for the manufacture of isoprene from glucose. The volatile isoprene is removed from the bioreactor with the waste gas, from which it can be condensed in high purity. Related procedures for the industrial production of isoprene have been reported by Ajinomoto, which cooperates with the Japanese tire producer Bridgestone, and by Michelin, which has chosen Amyris (USA) as a partner. The Malaysian state enterprise BioXCell cooperates with GlycosBio (USA).

1,3-Propandiol



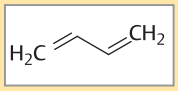
$C_3H_8O_2$	
M_R	76.10
D	10.5 g·cm^2 (20 °C)
b.p.	213 °C
CAS	504-63-2

i-Butanol, 2-Methyl-1-propanol



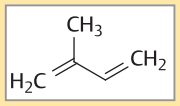
$C_4H_{10}O$	
M_R	74.12
D	0.80
b.p.	108 °C
CAS	78-83-1

1,3-Butadiene



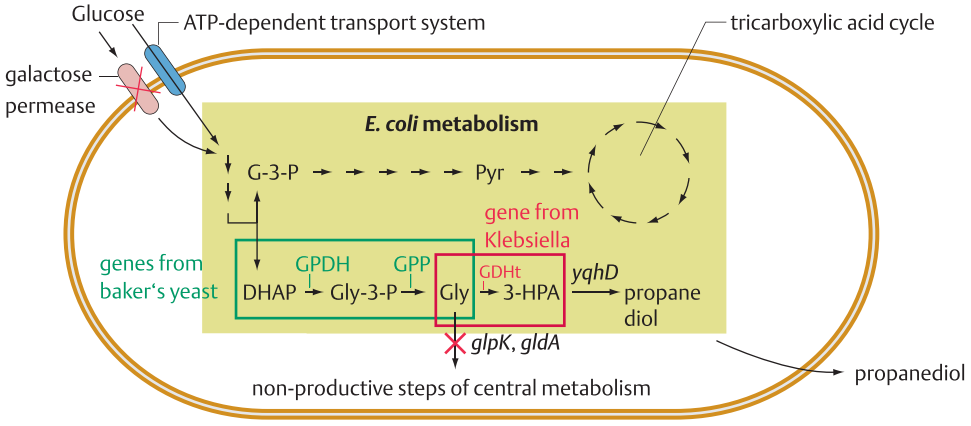
C_4H_6	
M_R	54.09
D	2,4982 (0 °C)
b.p.	−4.5 °C
CAS	106-99-0

Isoprene



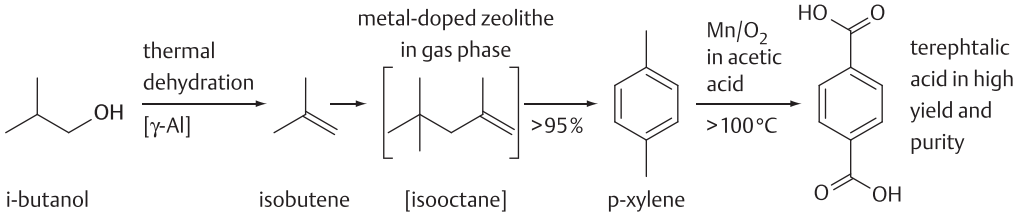
C_5H_8	
M_R	68.12
D	0.68
b.p.	34 °C
CAS	78-79-5

Manufacture of 1,3 propanediol



G-3-P	glyceraldehyde-3-phosphate
DHAP	dihydroxyacetone phosphate
Gly-3-P	glycerol-3-phosphate
Gly	glycerol
3-HPA	3-hydroxypropionic acid
GPDH	glycerol-3-phosphate dehydrogenase
GPP	glycerol-3-phosphate phosphatase
GDHt	glycerol dehydratase
yqhD	DHAP dehydrogenase
glpK	glycerol kinase
gldA	glycerol dehydrogenase
PTS	phosphotransferase transport system

Manufacture of terephtalic acid from i-butanol



Alcohols and alkenes which can be produced from renewable materials (selection)

C-2 intermediates	*C-3 intermediates*	*C-4 intermediates*	*C-5 intermediates*
ethene (ethylene) from bioethanol (Braskem, Brazil); ethane-1,2-diol (ethylene glycol) from cellulose by metal catalysis (CAS, China)	propanol, 2-propanol: through catalytic dimerization of bioethanol and metathesis, product is propene (propylene); 1,3 propanediol from glucose with engineered E. coli (DuPont)	2-butanol from glucose with engineered yeast, product is isobutene (Gevo); 2,3-butanediol from glucose with engineered Klebsiella or Enterobacter (Lanza) 1,4-butanediol from glucose with engineered E. coli (Genomatica); 1,3-butadiene through dehydration of 2,3-butanol, product is butadiene rubber (Lanxess)	Isoprene from glucose with engineered E. coli, product is polyisoprene rubber (Genencor, Ajinomoto, Amyris)
			C-15 intermediates
			β-farnesene, a sesquiterpene, from glucose with engineered yeast (Amyris, as a fuel)

Acetic acid/vinegar

General. Vinegar is used in many cultures for the acidification and preservation of vegetables, salads, rice, and other food products. Its consumption in these foods and in refreshing drinks is documented back to antiquity. It was and still is produced from fermented fruit juices, e. g., wine. In the 18th century, an "immobilization procedure" was developed in France whereby diluted wine was trickled over twiglets contaminated with acetic acid bacteria. Louis Pasteur succeeded in 1868 in defining selective growth conditions for acetic acid bacteria, thus laying a foundation for the modern technological production of vinegar. These days, vinegar is produced from ethanol by fermentation using strains of *Acetobacter*. If wine is used as a base material, the product is wine vinegar (a 6% solution of acetic acid in water, pH ~4.8); if rectified ethanol is used, the concentration of vinegar is 5%. Annual production in the USA alone is ca. 750 million L or 750,000 t. Glacial acetic acid (99.7%) is an important base chemical. It is produced from ethylene by catalytic oxidation and has a pK_a of 5.6. In the US, calcium magnesium acetate (m. p. –7.7 °C) (Cryotech CMA™) is used for deicing runways as it is less corrosive compared to sodium chloride. CMA produced from cornstarch has been proposed as a "green" antifreeze ("Nicer De-Icer").

Organisms and biosynthesis. Only a few species of Gluconobacter and Acetobacter can oxidize acetic acid to ethanol by "subterminal oxidation". Taxonomic classification of these species is complicated due to a rapidly changing phenotype during growth and is usually carried out by typing the 16S RNA, more recently also by analysis of plasmid profiles. Oxidation of ethanol proceeds via a sequential reaction of alcohol dehydrogenase (ADH) and aldehyde dehydrogenase (ALDH), which are both membrane-bound enzymes that contain pyrolloquinoline quinone (PAA) as prosthetic groups. ADH contains an additional heme C residue. They transfer the electrons generated by ethanol oxidation to a membrane-bound terminal oxidase via ubiquinone. During growth, these bacteria metabolize glucose to pyruvate both through glycolysis and through the KDPG pathway and further through the citric acid cycle. Both strains are extremely sensitive to lack of O_2. An interruption of the oxygen supply for even a few minutes results in a significant decrease in ethanol oxidation. If ethanol is depleted, acetic acid in the presence of O_2 is further oxidized to CO_2.

Fermentation and recovery. For the technical production of acetic acid, *Acetobacter* sp. is employed. This microbe is cultivated in a mash of aqueous wine or rectified ethanol, other nutrients and >60 g L^{-1} acetic acid under strong aeration, to prevent further oxidation of acetic acid. The process is carried out in a repeated fed-batch mode (→192): once the ethanol concentration has decreased to about 0.2% (ethanol sensor), a certain amount of the fermenter broth is removed and replaced with fresh mash. Since very homogenous aeration is necessary (0.1vvm, 0.1 volumes of air per fermenter volume min^{-1}), a highly efficient rotor/stator stirrer with self-priming aerator is used ("Frings Aerator"). Acid formation begins rapidly, with the formation of heat that is removed by exchangers. The average productivity of a 100 m^3 bioreactor using this process is ca. 1.6 g acetic acid $L^{-1} h^{-1}$. With special starter cultures and suitable monitoring and control, this process leads to a ~17.5% vinegar solution in 50–70 h. A more concentrated solution (up to 21%), as required in the canning industry, can be obtained if fermentation is continued for 45–55 h. Once a concentration of ca. 20% acetic acid has been reached, the acetic acid bacteria die and fermentation comes to an end. The raw vinegar is filtered and purified by a membrane process, pasteurized, and diluted to a 5–6% vinegar solution which can be marketed. Among the popular designs for vinegar production is the Frings Acetator. Other process variants, e. g., continuous fermentation with cell recycling or the use of immobilized acetic acid bacteria in an airlift bioreactor (→96), sometimes show higher productivity (up to >100 $gL^{-1} h^{-1}$) but have not yet found wide acceptance in the marketplace. The manufacture of traditional vinegars proceeds through a slow fermentation of selected musts of many regional variations (e. g., grape, wines, coconut water) over weeks or months. The desired aroma develops during maturation, which may take years, as in the traditional Balsamic vinegars of Modena, Italy. Usually, the Acetobacter remain as slime in the vinegar.

Acetic acid

$C_2H_4O_2$		
M_R	60.05	chemical synthesis:
Sdp.	117.9 °C	
pK_a	4.76 (25 °C)	addition of O_2 to ethylene
CAS	64-19-7	or of CO to methanol

$H_3C-COOH$

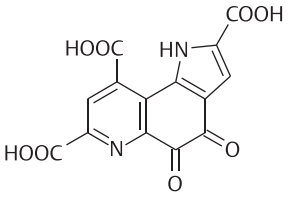
Biosynthesis by *Acetobacter sp.*

alcohol dehydrogenase: $C_2H_5OH \rightarrow H_3C-CHO$ (acetaldehyde)

aldehyde dehydrogenase: $H_3C-CHO \rightarrow H_3C-COOH$ (acetic acid)

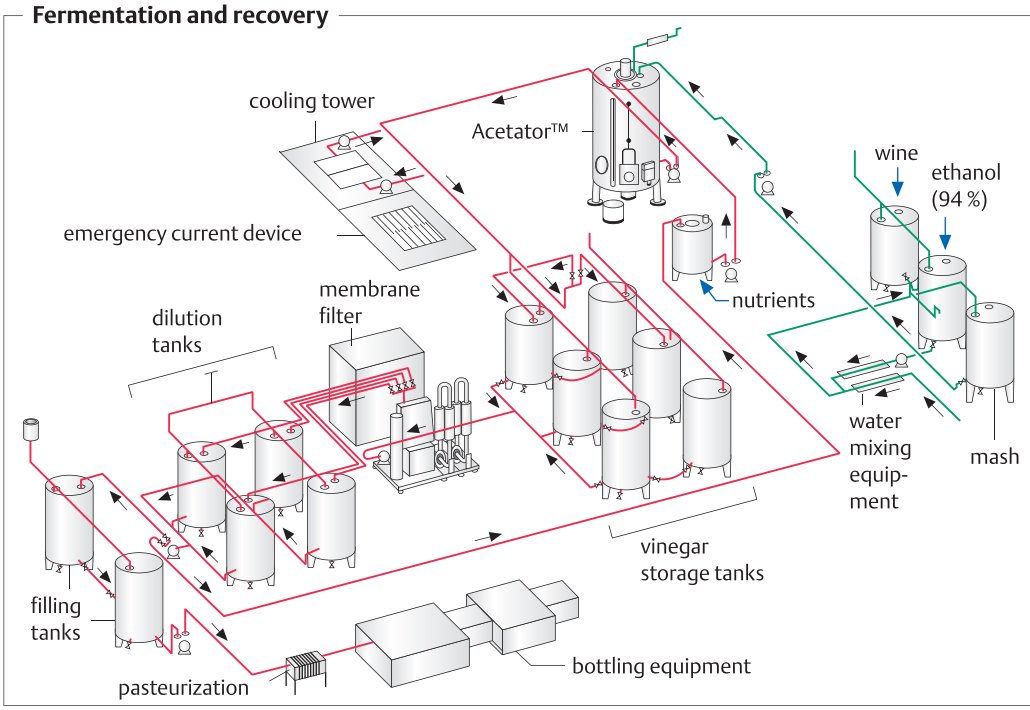
$H_2O + CO_2$ ← "subterminal oxidation"

the membrane-bound, PQQ-dependent dehydrogenases transfer the electrons generated by the oxidation of ethanol *via* ubiquinone to a membrane-bound terminal oxidase



pyrroloquinoline quinone (PQQ)

Fermentation and recovery



Process alternatives

	maximum vinegar production [%]	productivity [L/m³ d]	remarks
standard procedure (repeated batch)	15	35 – 50	simple process scheme
one-step high-percentage procedure	18.5	30 – 50	high vinegar concentration, low storage and transportation cost
two-step high-percentage procedure	> 20	30 – 50	high vinegar concentration, low storage and transportation cost
continuous process	> 10	up to 60	high vinegar concentration, low storage and transportation cost
immobilized acetic acid bacteria (experimental)	< 9	–	fluid bed- or airlift reactors, up to 460 d

Citric acid

General. Citric acid was first isolated in 1822 from lemon juice by Carl Wilhelm Scheele, who also established its composition. Many fruits form large quantities of citric acid. In 1934, Hans Krebs discovered that citric acid is a central compound in aerobic metabolism (citric acid cycle) (→26); for example, in the metabolism of an adult human, 1.5 kg citric acid are formed daily as an intermediary product. Citric acid is a strong tribasic acid. The pK_a-values of its three dissociation steps are 3.13, 4.78, and 6.43 (25 °C). A 1% solution of citric acid in water has a pH of 2.2. With its three carboxyl and one hydroxyl group, citric acid is an excellent complexing agent for di- and trivalent cations. Citric acid is exclusively produced by fermentation. Over 2/3 of the 2 million tons produced annually (2012) are made in China, and the market value is ca. 1.8 billion USD. Citric acid is used as an acidulant and preservative in the food industry, as a complexing agent in metal treatment, as a water softener in detergents and as a therapeutic agent for heavy metal poisoning in emergency medicine.

Organisms and biosynthesis. Some molds, such as *Aspergillus niger* (→16), secrete large quantities of citric acid during and after the late logarithmic growth phase, provided there is an excess of glucose and oxygen. Although the intermediates of the citric acid cycle are usually consumed by the general metabolism, the nearly quantitative conversion of glucose to citric acid by *A. niger* is possible for two reasons: 1) oxaloacetic acid, an intermediate of the citric acid cycle, is replenished via an anaplerotic reaction and 2) the enzyme pyruvate carboxylase, which is localized in the cytoplasm in this mold, catalyzes the addition of CO_2 to pyruvate and thus forms oxaloacetic acid by a shortcut from glycolysis. In addition, citric acid is secreted from the mitochondria, where it is formed, into the cytoplasm by an antiporter which in turn imports maleic acid (a reduction product of oxaloacetic acid) from the cytoplasm. The rate-determining step for citric acid biosynthesis is the transport of glucose followed by phosphorylation.

Manufacture process. *A. niger* is used for the industrial production of citric acid from sugars. A small part of the production is still carried out by a traditional surface fermentation process (→86): acid-resistant trays in a sterile compartment are filled with the sugar mash and inoculated with spores of *A. niger*. After 5 days, a mycelial mat has formed on the surface that performs the fermentation. A high degree of aeration (up to 10vvm, volumes air/volume broth/min) is required, mainly to remove the generated heat. After ca. 8 d of surface fermentation, the mycelium can be removed, extracted with hot water, and citric acid is precipitated from the combined liquids. Yields are on the order of 50 g kg^{-1} sugar. Today, most citric acid is produced in stirred or airlift fermenters (→96) of 100–500m^3 volume. The stainless steel reactors (harvest solution: pH 2.0!) are first sterilized by steam and then filled with an inexpensive carbon source (→88) such as starch hydrolysate or sucrose. By a mechanism that is not fully understood, high citrate concentrations are favored in media limited in Mn^{2+} (< 2 μg L^{-1}, obtained through the addition of hexacyanoferrate or cation exchangers). Under these conditions, the mycelium does not grow out to a loose mycelial mat with few spores but rather forms small, solid pellets with a diameter < 0.5 mm. Formation of the cell mass is usually complete after 48 h at pH 5. The addition of sugar in a fed-batch mode and an increase of aeration initiate formation of citric acid, which is excreted into the medium. Yields are in the order of 150 g/L and > 80% molar yield. Traditionally, the mycelium was removed by filtration, citric acid was precipitated from the filtrate by addition of $Ca(OH)_2$, and recovered from the resulting calcium citrate by reacting with sulfuric acid. Addition of activated charcoal or ion exchangers allowed for crystallization of very pure citric acid. During the above process, > 1 t gypsum/t citric acid was formed, resulting in high waste and wastewater treatment costs (→108). Recovery is mostly carried out by sequestering citric acid in the fermentation broth with trilauryl amine and then extracting the complex with a mixture of alkanes and 1-octanol. Solvents and sequestering agents can be recovered in this process. Some alkane-degrading yeasts such as *Candida parapsilosis* are able to form citric acid from the less volatile alkane fractions. However, this technology has been abandoned due to rising oil prices and concerns about potentially hazardous residues from petroleum components remaining in a food product.

Citric acid

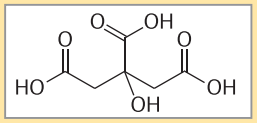
$C_6H_8O_7$ CAS 77-92-9

M_R 192.12

solubility 600 g/L water (20 °C)

acidity pK_{a_1} = 3.128 (25 °C); 1 % solution of citric acid in water: pH 2.2

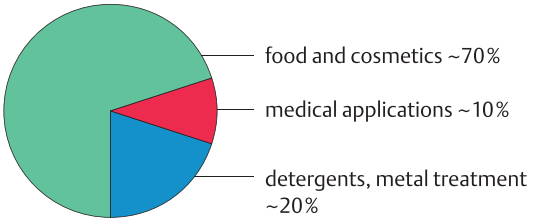
complexing constants (lgK): Fe^{3+} 12.5, Ca^{2+} 4.68, Cu^{2+} 3.98 (at 20 °C)



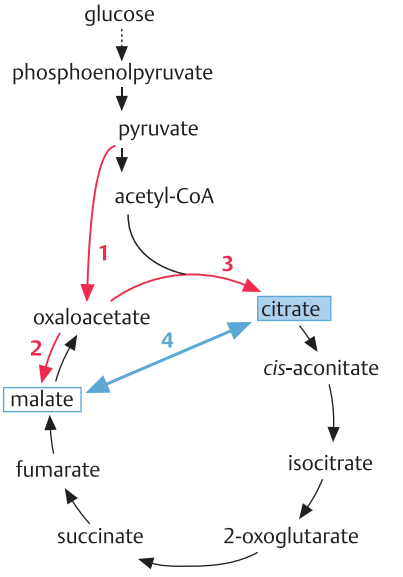
Occurrence and use

occurrence	[g kg^{-1}]
lemons	40 – 80
grapefruit	12 – 21
raspberries	10 – 13
black currants	15 – 30
strawberries	6 – 8
tomatoes	2.5

World market: ca. 2 million to (2012)

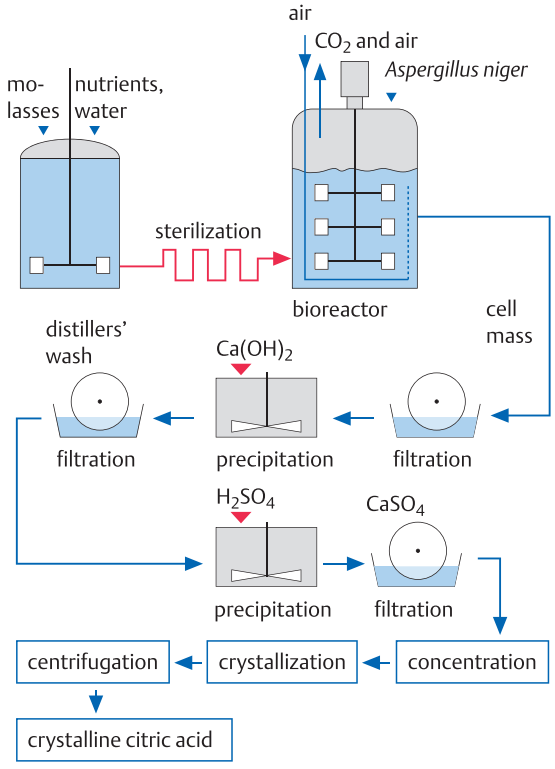


Biosynthesis in *Aspergillus niger*



anaplerotic reactions of the citric acid cycle:

1 pyruvate carboxylase (in cytoplasm)
2 malate dehydrogenase (in cytoplasm)
3 citrate synthase (in mitochondria)
4 citrate/malate antiporter (mitochondrial membrane)



Fermentation and recovery

inoculation material, medium	seed fermenter	bioreactor	filtration	purification	
starch or molasses, N-source, salts	40 m^3 bio-reactor	400 m^3 stainless steel. 8 d at 32 °C. After cell growth pH shift from 5.0 to 2.0	removal of mycelium by vacuum rotary filter	precipitation, recrystallization 1. $Ca(OH)_2$ 2. H_2SO_4 3. recrystallization	> 200 g/L citric acid after 150 h

Lactic acid, 3-hydroxy-propionic acid (3-HP)

General. In 2013, about 300,000t of lactic acid (→116) were produced, mostly by fermentation. A major application of this chemical is in food and beverages, due to its pleasant acidic taste and its preservative properties. A less pure product is used in the leather and textile industries. A fast growing application of both D- and L-lactic acid is as a chemical building block in the synthesis of biodegradable polyesters (NatureWorks™) (→154). 3-hydroxypropionic acid (3-HP) is a non-natural hydroxy acid that can be produced from metabolically engineered (→318) *E. coli*. Like lactic acid, it is an attractive building block for polymers, but also a platform chemical for the synthesis of other chemicals such as acrylic acid (obtained from 3-HP by dehydration). Polymers based on acrylic acid are used in diapers, coatings, textiles, adhesives, water treatment, etc.

Lactic acid: organisms and biosynthesis. L-lactic acid is produced technically by various Lactobacilli. The choice of organism depends on which carbon source is used. For an exhaustive transformation of the substrate, homofermentative Lactobacilli (→116) must be used, since they produce two moles of L-lactic acid per mole of D-glucose during glycolysis. D-lactic acid is produced by, e. g., *Sporolactobacillus laevolacticus*. Fungi and bakers' yeast can also be metabolically engineered (→318) to overproduce lactic acid. Space-time yields, however, are presently not competitive to lactobacilli-based processes.

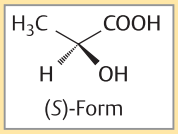
Fermentation and recovery. Lactic acid can be produced via chemical synthesis or fermentation. Chemical synthesis is done by the addition of H_2O to acrylic acid, or of HCN to acetaldehyde, and leads to racemic lactic acid. In fermentation, the choice of organism depends on the carbon source. *Lactobacillus delbrueckii* or *L. leichmannii* is preferred if dextrose or other sugar solutions are used, and *L. bulgaricus* is used on whey as a carbon source (→88). The fermentation medium contains 12–18 % sugar, a nitrogen source, phosphate, and B vitamins. Because lactobacilli lose activity below about pH 4.5, fermentation processes are operated at pH values between 5.5 and 6.0, at 45–50 °C under O_2-poor conditions in the presence of $CaCO_3$ as a buffer (to keep the pH constant between 5.5 and 6.0) and are complete after 2–6 days, depending on the substrate concentration. After removal of the cell mass, Ca-lactate can be transformed by the addition of H_2SO_4 into the free acid, which can be further purified by ion exchange chromatography. Alternatively, esterification with methanol yields lactic acid methyl ester, which can be purified by distillation. The use of liquid membranes and the direct use of ion exchangers in the fermentation broth, without prior precipitation of the Ca salt, are under development. For polyester applications, a lactide is formed through condensation and purified by vacuum distillation.

3-Hydroxypropionic acid. In order to produce this non-natural compound in a microorganism from glucose or glycerol as a raw material, several artificial metabolic pathways (→320) were constructed and investigated. A preferred pathway uses an *E. coli* host and starts from glucose. It leads via pyruvate and L-alanine to ß-alanine, which can be transformed in two enzymatic steps through malonyl semialdehyde to 3-HP. This pathway is of interest insofar as an enzyme which isomerizes L- to ß-alanine (alanine-2,3-aminomutase) had not yet been found in nature. As a consequence, this enzyme activity was engineered into a related protein, lysine-2,3-aminomutase, by a combination of protein design (→198) and enzyme evolution and found to be active with L-alanine after expression in *E. coli*. Industrial optimization of this *E. coli*-based process has meanwhile reached a pilot scale (2014). Alternative concepts use, e. g., glycerol as a C-source, which is dehydrated to 3-hydroxy-propionaldehyde, using a gene for glycerol dehydratase from *Klebsiella pneumonia*, which is further reduced to 3-HP by an α-ketoglutaric semialdehyde dehydrogenase obtained *from Azospirillum brasiliense*. As this process needs the addition of vitamin B12 to the medium, and *K. pneumoniae* forms B12 in its cytoplasm, it has been attempted to express the whole pathway in *K. pneumoniae* as a host organism. The yields, however, are still unsatisfactory and in the range of 40g/L. The above developments are examples for a new "synthetic biology" (→320) which is no longer based on natural pathways and enzymes, but attempts to design both the pathway and the required enzymes at will.

D-Lactic acid, L-Lactic acid

$C_3H_6O_3$
M_R 90.08
pK_a 3.80 (25 °C)
CAS 50-21-5
10326-41-7 (*R*)-Form
79-33-4 (*S*)-Form

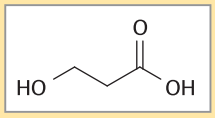
chemical synthesis of racemate:
addition of HCN to acetaldehyde



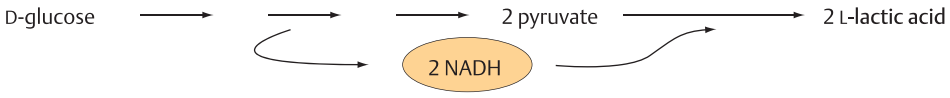
(*S*)-Form

3-Hydroxypropionic acid

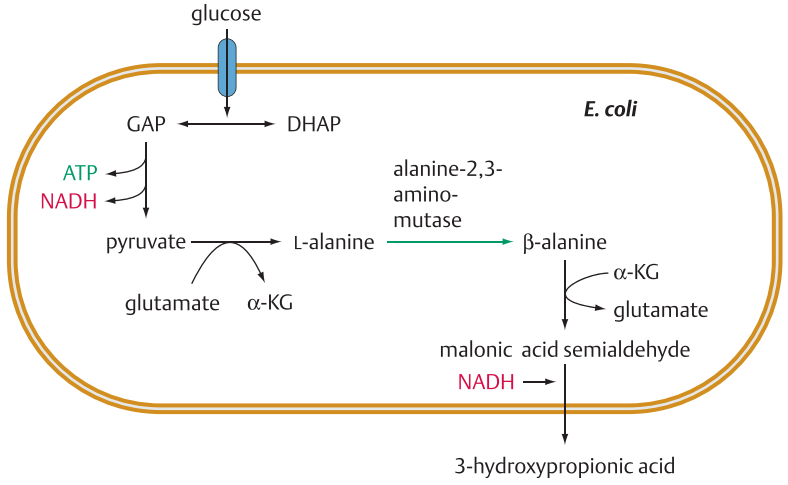
$C_3H_6O_3$
M_R 90.8
pK_a 4.50 (25 °C)
fp 213 °C
CAS 503-66-2



Biosynthesis of L-lactic acid (example: *Lactobacillus delbrueckii*)

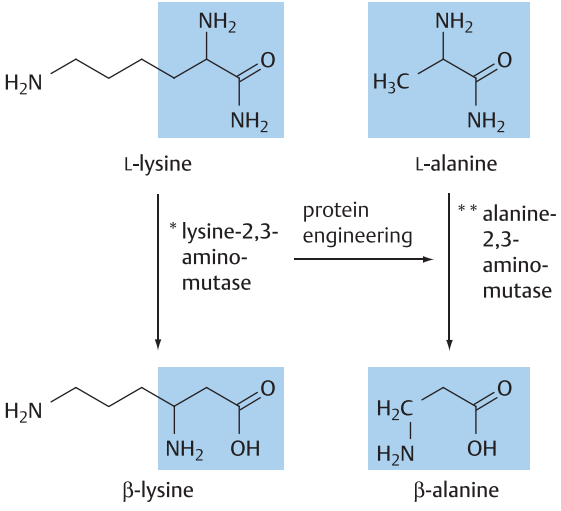


Synthesis of 3-hydroxypropionic acid by *E. coli* through the β-alanine pathway



α-KG α-keto-glutarate
GAP glyceralde-hyde phosphate
DHAP dihydroxy-acetone phosphate

"Synthetic" alanine-2,3-aminomutase



*naturally occuring enzyme
** not naturally occuring enzyme

Fermentation and recovery

	L-lactic acid	3-hydroxy-propionic acid
preculture	Lactobacillus	*E. coli* recombinant
bioreactor	> 100 m^3, dextrose or sugar solutions, 33 °C, pH 6.5, 1 vvm	glycerol or glucose
recovery	removal of cells, precipitation as Ca salt, dissolution, ion exchange	removal of cells, acid extraction, precipita-tion
	productivity 2 – 3 kg m^{-3} h^{-1}	target: >90% of theoretical yield based on glucose

Gluconic acid and "green" sugar chemicals

General. Gluconic acid, produced by selective oxidation of D-glucose in a fermentation process, is an established product of biotechnology. More recently, sugar derivatives have been developed in the framework of a "green technology". Examples are 2,5-furandicarboxylic acid, levulinic acid esters, glucaric acid and isosorbide. Produced by chemistry, they are potential building blocks that might become available from future "biorefineries" (→320).

Gluconates. Na-D-gluconate, D-gluconic acid, and its δ-lactone are produced at a level of ca. 70,000t. The δ-lactone is used in the food industry as a mild acidulant. Ca^{2+}- and Fe^{3+}-gluconates are highly soluble and nontoxic and thus are used in medical infusion preparations for the treatment of calcium or iron deficiencies. Sodium gluconate is also a highly alkali-stable complexing agent for calcium and iron; ca. 50 % of its production is used as an additive for bottle cleaning and alkaline de-rusting, in the preparation of concrete, and for the prevention of iron precipitates in textile treatments. The pK_a of D-gluconic acid is 3.7. D-gluconic acid is the end product of a subterminal oxidation of D-glucose and thus similar to the subterminal oxidation of ethanol to acetic acid. Some molds (*Aspergillus niger, Penicillium* species, →16), and also oxidative bacteria, especially Gluconobacter, carry out this reaction. In molds, the responsible flavoenzyme is D-glucose oxidase, which is localized in the fungal cell wall but can also be found in the medium during fermentation. Glucose oxidase is a key enzyme for the determination of blood sugar levels in biosensors. In contrast, Gluconobacter strains use a membrane-bound D-glucose dehydrogenase that contains pyrolloquinoline quinone as a cofactor, similar to the alcohol and aldehyde dehydrogenases of Acetobacter strains.

Fermentation and recovery. D-Gluconic acid is prepared from D-glucose by electrochemical oxidation or by a fermentation process using *Aspergillus niger.* At pH-values > 3, this mold accumulates glucose oxidase in its cell wall, which oxidizes D-glucose to D-glucono-5-lactone, which hydrolyzes spontaneously or faster by enzymatic catalysis (lactonase) to D-gluconic acid. Na- or Ca-gluconate is obtained upon growth of the cell mass at pH 4.5–6.5 (buffered with Na_2CO_3/NaOH or $CaCO_3$) by the addition of 11–25 % D-glucose under strong aeration. The salt is obtained from the filtered fermentation broth by concentration and drying. The free acid and the lactone are obtained from the salt by ion-exchange chromatography.

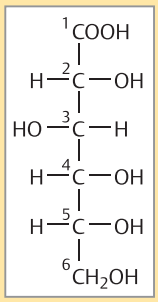
2,5-Furandicarboxylic acid (FDCA) is a human metabolite and can be detected in urine or blood plasma. It is technically prepared from D-glucose by dehydration to hydroxymethylfurfural or alkoxymethylfurfural, followed by catalytic oxidation under strongly alkaline conditions. The oxidation step is also selectively done by a bacterium, *Cupriavidus basilensis* HMF14. FDCA is a potential replacement for petrochemistry-based terephthalic acid and can be condensed with diols such as ethylene glycol (e. g., from bioethanol) to "green" materials such as polyethylene furanoates (PEF), which are undergoing testing as a replacement to polyethylene terephthalate (PET); PEF is completely made from sugar and is biodegradable. This so-called Y-X-Y ("iksy") technology is promoted by Avantium, a Dutch company that cooperates with several bottling companies such as Coca Cola or Danone. By exchanging the diol component, e. g. to propanediol, other materials such as fibers or films can be produced using the Y-X-Y technology.

Isosorbide is a very hygroscopic material, obtained by two dehydration steps from D-sorbitol. It is used as a desiccant and diuretic but also has potential as a chemical building block, e. g., for the synthesis of polyesters. Isosorbide polycarbonate (Durabio™) has similar properties like petrochemically produced polycarbonates.

Glucaric acid is obtained by selective chemical oxidation of D-glucose. The dicarboxylic acid contains four chiral centers and thus allows for the preparation of numerous chiral products. Glucaric acid is already used as a mild cosmetic exfoliant.

Levulinic acid esters are prepared by heating D-glucose in diluted acids in the presence of alcohols. In the presence of methanol, for example, levulinic acid methyl ester is formed; in the presence of glycerol levulinic acid ketals are formed. Such products have a potential to replace petrochemical reagents, e. g., in the fields of plasticizers, solvents, polyurethane components or agrochemicals such as δ-aminolevulinic acid.

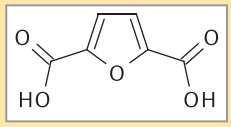
D-Gluconic acid



$C_6H_{12}O_7$

M_R	196.16
pK_a	3.7 (25 °C)
CAS	526-95-4

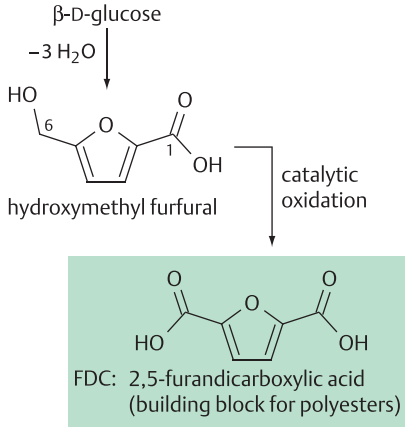
2,5-Furandicarboxylic acid



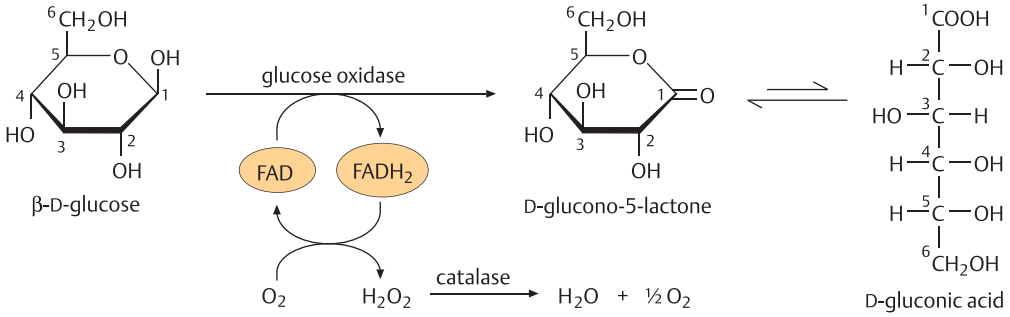
$C_6H_4O_5$

M_R	156.09
pK_a	2.28
CAS	3238-40-2

Synthesis of FDC



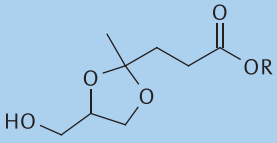
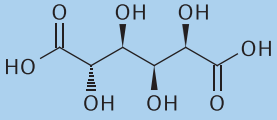
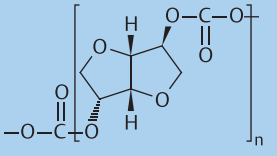
D-gluconic acid (*Aspergillus niger*)



Fermentation and recovery of D-Gluconic acid

preculture	bioreactor	recovery	
Aspergillus niger	> 100 m^3, dextrose or sugar solutions, 33 °C, pH 6.5, 1 vvm	cell separation, concentration, addition of NaOH, drying	productivity > 13 kg m^{-3} h^{-1}

Other "green" sugar building blocks

Product	Structure	Manufacture	Applications	Status	Company
levulinic acid ketal ester		thermochemically from glucose and pentoses	plasticizer	pilot plant	Segetis, USA Aventium, NL
glucaric acid		chemical oxidation of glucose biopolymers	biopolymers, polyamides	pilot plant	Rivertop Renewables, USA
isosorbide (polycarbonate)		from sorbitol by chemical dehydration	biopolycarbonate	production, Durabio™	Mitsubishi Chemical, Japan Roquette, France

Dicarboxylic acids

Dicarboxylic acids are important industrial building blocks for polyesters (→154), polyamides (→156) and other products. Most dicarboxylic acids are chemically synthesized through the oxidation of alkanes or cycloalkenes. An exception is sebacic acid (octane-1,8--dicarboylic acid), which is formed by alkaline cleavage of ricinoleic acid (9*Z*,12*R*)-12-hydroxy-9-octadecenoic acid), a component of castor oil. Some dicarboxylic acids can be produced by fermentation. This field is stimulated by a growing need for a diversity of dicarboxylic acids for the synthesis of biopolymers.

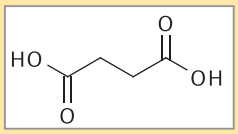
Succinic acid is an intermediate of both the citric acid and the urea cycle (arginine succinate). It is a food-grade acidulant, but also an important platform chemical for the synthesis of biopolymers such as polyesters, polyamides, and polyfuranes. For chemical synthesis, C-4 diacids such as fumaric acid are used. For industrial fermentation, mutants of *E. coli*, baker's yeast, the Gram-negative rumen bacterium *Basfia succinoprodugenes* or other strains are used. The *E. coli* mutant has been engineered to suppress the formation of undesired side products such as acetic acid, formate or lactic acid through deletion of the gene for pyruvate-formate lyase and lactate dehydrogenase. Upon incorporation of a gene coding for pyruvate carboxylase, cloned from *Rhizobium etlii,* the tricarboxylic acid cycle is filled up by an anaplerotic sequence. The industrially used bakers' yeast strain produces succinic acid at lower pH values as compared to *E. coli*, facilitating pH correction during the fermentation and leading to lower salt loads in recovery. Succinate yields were enhanced 30-fold relative to the wild type strain through deletion of the gene *sdh3*, which couples formation of succinic acid with the generation of biomass, or deletion of *ser3/ser33,* which channels 3-phosphoglycerate into the serine metabolism.

Itaconic acid is a side product of the citric acid cycle that is formed by the fermentation of sugar by various Aspergillus strains, e. g. by *Aspergillus terreus*. The biosynthesis of itaconic acid proceeds *via* citric acid, which is first dehydrated to *cis*-aconitic acid and then decarboxylated to itaconic acid. The fermentation follows the procedures of citric acid production, and yields are on the order of 80 g/L. Recovery is by crystallization of a concentrated itaconic acid solution at 15 °C. Itaconic acid is chemically related to methacrylic acid but contains one more functional group. It is a useful building block for the synthesis of many polyacrylates, resins and surfactants. In 2009, 50,000t of itaconic acid were produced.

Higher dicarboxylic acids. Yeasts of the genus Yarrowia, e. g., *Yarrowia lipolytica* or Candida, e. g., *Candida tropicalis* (→14), grow on alkanes as a sole carbon source. In the metabolism of alkanes, the first step is a monoterminal, diterminal or subterminal oxidation of the alkane, employing monooxygenases. The resulting hydroxylated compounds are further oxidized by dehydrogenases to carboxylic acids, which undergo assimilation of the general metabolism through ß-oxidation to acetyl-CoA. Fatty acids can also be oxidized to dicarboxylic acids by monoterminal oxidation and dehydrogenation. When ß-oxidation is blocked, yeast mutants can be obtained which transform alkanes or fatty acids into dicarboxylic acids or ω-hydroxy fatty acids. Fatty acids which are renewable materials obtained mostly from plant oils thus allow for the synthesis of dicarboxylic acids as chemical building block. The disadvantage of this process is, however, that fatty acids with a chain length below C-12 (lauric acid) are not available in industrial quantities, limiting the preparation of dicarboxylic acids by this procedure to chain length of C-12 to C-18. However, polyamide fibers, the largest product among the polyamides, require shorter dicarboxylic acids such as

Adipic acid is a product rarely encountered in nature. It is chemically produced by oxidation of cyclohexene or cyclohexenone. Adipic acid (hexane dicarboxylic acid) is a component of nylon 6,6 (polyester from adipic acid and hexamethylene diamine). Recently, recombinant yeasts were metabolically engineered (→318) to produce adipic acid from alkanes, fatty acids or glucose (Verdezyne, USA). In these strains, an acyl-CoA synthase enzyme in the fatty acid degradation system of the host (Candida tropicalis) is substituted by a hexanoate-CoA synthase from *Aspergillus parasiticus*. In addition, genes for unnecessary branched pathways are deleted. Adipic, sebacic and dodecanedioic acid are promising future building blocks for "green" polyesters and polyamides.

Succinic acid

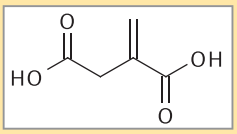


$C_4H_5O_4$

M_R	118.09
pK_s	4.16 and 5.51
fp	184 °C
solubility	in water: 58 g·l^{-1} (20 °C)
CAS	110-15-6

chemical synthesis: catalytic hydrogenation of maleic acid anhydride (example)

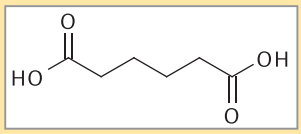
Itaconic acid



$C_5H_6O_4$

M_R	130.10
pK_s	3.84 and 5.55
fp	184 °C
solubility	in water: 63 g·l^{-1} (20 °C)
CAS	97-65-4

Adipic acid

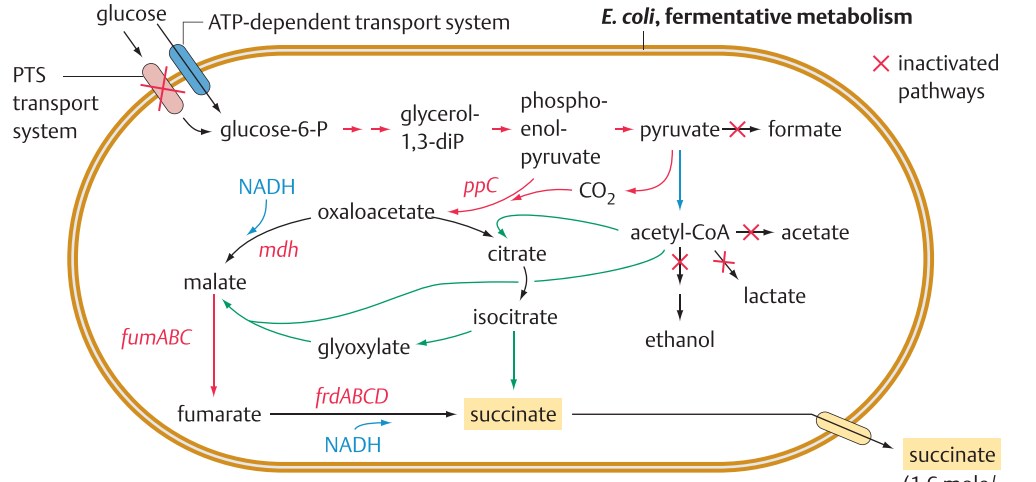


$C_6H_{10}O_4$

M_R	146.14
pK_s	4.43 and 5.42
fp	151 °C
solubility	in water: 24 g·l^{-1} (20 °C)
CAS	124-04-9

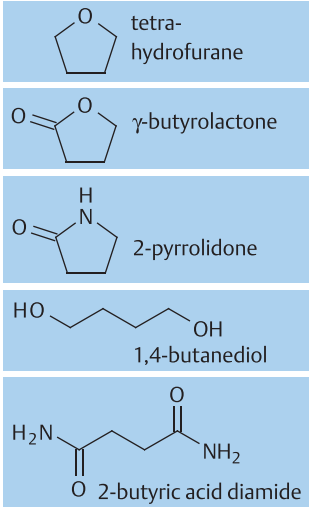
chemical synthesis: catalytic oxidation of cyclohexane

Biosynthesis of succinic acid

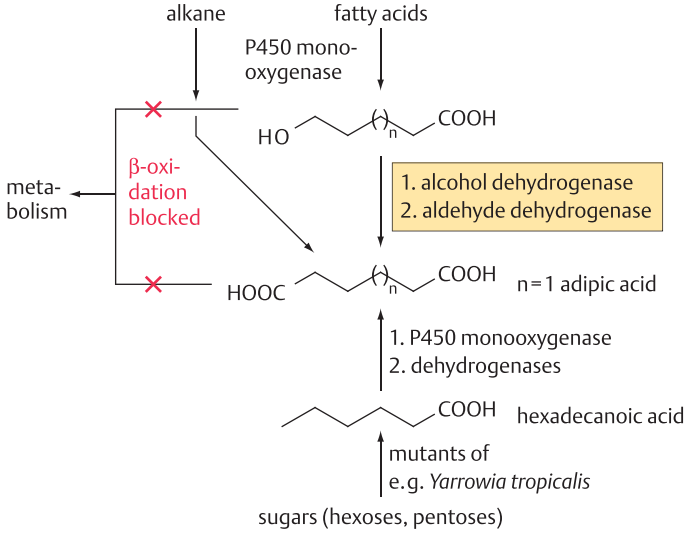


ppC gene for phosphoenolpyruvate carboxylase
mdh gene for malate dehadrogenase
fumABC gene for fumarase
frdABCD gene for fumarate reductase

Succinate derivatives



Dicarboxylic acids



Biopolymers: Polyesters

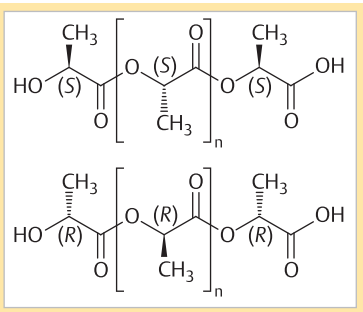
General. Polyesters are a large group of polymers that are produced by polycondensation of acids and alcohols obtained from petrochemical materials. Examples are polyethylene terephtalate (PET), which is used to produce, e. g., plastic sheets or bottles, and polycarbonate (PC), which is used for the manufacture of CDs, airplane windows and many other purposes. Biopolymers ("bioplastics") are a recent development to produce materials from sustainable resources such as starch or plant oils. Some biopolymers are more biodegradable (or decay much faster) compared to traditional plastics, since chiral building blocks that are substrates for enzymes of environmental microorganisms are used. Major examples of polyester-type biopolymers are polyacetate (PLA), the polyhydroxy-alkanoates (PHA) and "Bio-PET"-products. Polycarbonates based on isosorbide (→150) are also studied. Manufacturing costs for polyester biopolymers are higher as compared to traditional polyesters (2013), and their manufacturing and application properties need to be improved. Cheaper, and already more successful in the markets, is "thermoplastic starch" or "PSM" (for "plastarch material," or starch treated with plasticizers such as sorbitol and glycerol). Only some 8 % of all polymers (about 366 million t in 2011) are presently bio-based, but the annual growth of this product group is nearly 20 %.

Polylactides (PLA) have been industrially produced for many years from L-lactic acid (NatureWorks™) (→148). The fermentation product is often isolated as a dimer and transformed into PLA by ring-opening polymerization under metal catalysis. After compounding with additives, the material is extruded or cast into various products such as compost bags, food packaging items, or disposable tableware. For applications in the car or machine manufacturing industries, more demanding material properties are needed (e. g., higher melting temperature, higher impact resistance). Such features can be obtained by using a 1:1 mixture of L- and D-lactide for copolymerization ("stereo-complex" sc-PLA, m. p. ~230 °C). As a consequence, improved fermentation procedures for the industrial production of D-lactic acid have also been developed.

Polyhydroxyalkanoates (PHA). Many microorganisms, e. g., *Ralstonia eutropha*, can form, under appropriate conditions, poly-(*R*)-3--hydroxybutyric acid) (PHA) in quantities up to 90 % of their dry cell mass. In some strains, the composition of this polymer can be modified over a wide range by the addition of precursors such as lactic acid. The PHA operon of *Ralstonia eutropha* consists of only 3 genes and was successfully transformed into *Escherichia coli, Pseudomonas putida* and other microorganisms and plants. Of technical interest are copolymers of (*R*)-3-hydroxybutyric acid and (*R*)-3-hydroxyvaleric acid (Biopol®), which show properties similar to polypropylene but can be produced from renewable resources and are biodegradable. They are produced in small quantities for special applications in surgery. Another commercial product is Aonilex®, a copolymer from (*R*)-3-hydroxybutyric acid and (*R*)-3-hydroxyhexanoic acid (PHBH). This biopolymer reaches melting points up to 160 °C, depending on compounding, and has other good application properties. For its production, an engineered strain of *Cupriavidus necator*, a Gram-negative bacterium, is used. The carbon and nitrogen sources (→328) in fermentation are plant oil and inorganic nitrogen, and the yields based on available carbon are nearly quantitative. For recovery of PHA biopolymers, which are formed within the cytoplasm, enzymatic digestion of the cell wall, extraction by solvents or, simpler, drying and pelleting of the cells have been described.

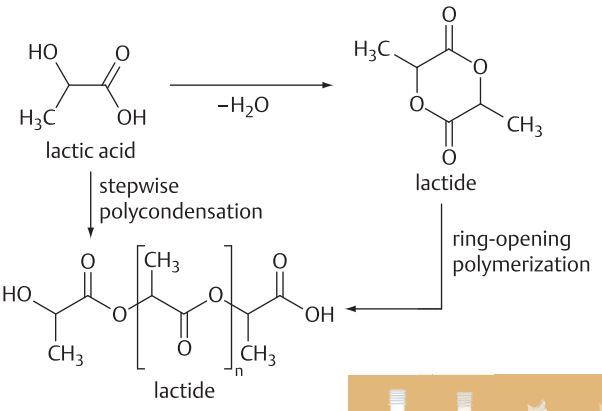
"Bio-PET" is a copolymer from diols prepared by fermentation with terephthalic acid (→142). A copolymer made from terephthalic acid and 1,3-propanediol produced by transgenic *E. coli* is already registered and used as Sorona® for the manufacture of textile fibers and carpets. A copolymer made from terephthalic acid and ethylene glycol, obtained by hydration of bioethanol, is being used for the manufacture of Coca Cola's "green bottle". In both cases, terephthalic acid is made from petrochemical raw materials; the percentage of biomaterial in the product is only around 30 %. Industrial projects are also underway to produce terephthalic acid from renewable resources, e. g., by dimerization of i-butanol produced by engineered yeast, or to replace terephthalic acid by 2,5-furandicarboxylic acid (FDC) (→150), which is synthesized by dehydration of glucose followed by chemical oxidation.

Polylactides



CAS-Nr. 33135-50-1

	m.p. [°C]	crystallinity [%]
Poly-L-lactide	170–180	37
Poly-D-lactide	170–180	37
scPLA (1:1 mixture)	230	>50



biocompostable bottle after 6 weeks in soil

Polyhydroxyalkoanoic acids

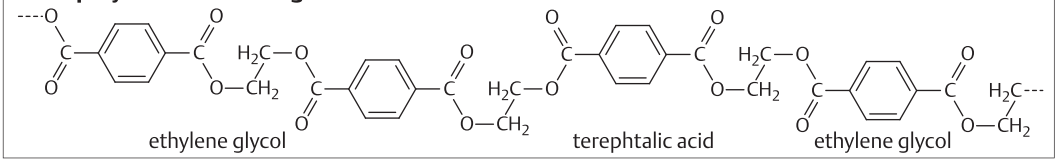
product	catalyst	manufacturer
copolymers of 3-hydroxybutyric acid and 3-hydroxyvaleric acid (Biopol®)	*Ralstonia eutropha*	Metabolix (commercial) (Mirel)
copolymers of 3-hydroxybutyric acid and 3-hydroxyhexanoic acid (Aonilex®)	*Ralstonia eutropha* gene cluster from *Aeromonas caviae*	Kaneka ("PHBH")
copolymer of butane-2,3-diol and succinic acid ("GS Pla®")	chemical Synthesis from 2 bio monomers	Mitsubishi Chemical
PHB, cloned operon from *Alcaligenes latus*	*Escherichia coli*	ATO-DLO
polyester with side-chains ranging from C_3- to C_8- or phenylvaleriate side-chain	*Pseudomonads*	not commercial
PHB, cloned operon from *Ralstonia eutropha*	rapeseed	not commercial

Ralstonia eutropha	**bioreactor**	**recovery**	
heterotrophic production strain	fed-batch, glucose, propionic acid, 30 °C, ~ 40 h	enzymatic break-up of cells	> 80 g/L culture medium, > 2 g/L × h

PHB operon from *R. eutropha*

promoter	PHA synthase	3-ketothiolase	acetyl-CoA reductase

PET polymers and analogs



PET polymers and analogs

diol component	acid component	product/trade name	company
glycol from bioethanol	terephtalic acid	"green bottle"	Coca Cola
glycol from bioethanol	terephtalic acid from isobutanol	under development	Gevo
glycol from bioethanol	furane dicarboxylic acid	under development	Aventium
1,3-propanediol	terephtalic acid	Sorona textiles	Dupont

Biopolymers: Polyamides

General. Peptides, proteins and enzymes are, by their structure, polyamides. Silk, also a polyamide, has been used for thousands of years as a textile material in Asia. Some other polyamides have a potential as "smart polymers" and can be produced by biotechnology. Among these are the dragline silk of spiders, the adhesion proteins of mussels or the composite materials of nacre ("molecular bionics"). Important classes of polyamides are at present produced from petrochemical base materials, e. g., nylon® and perlon®. They serve as thermoplastic fibers, cords, anchor bolts or insulation materials. While the production of such polyamides from renewable materials ("Bio-Nylon") also has considerable biotechnological potential, the raw materials are mostly too expensive yet for a broad use.

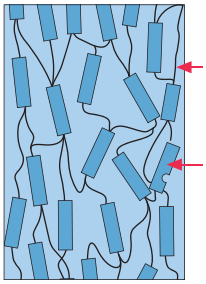
Silk consists of fibroin as a structural component. Fibroin is built from repeating units of hexapeptides and spacers. The high proportion of glycine and the formation of many hydrogen bonds lead to a very tight and stable packaging of β-sheets (→28). Fibroin is produced by the caterpillar of the silk moth *Bombyx mori* and secreted together with the auxiliary protein sericine. The high capacity of this secretory system was already used for the production of transgenic proteins such as erythropoietin, ferritin or spider silk, and such processes are on the verge of industrial use in Japan (2014).

Spidroins are highly variant polyamides extruded by the spinnerets of spiders (e. g., *Nephila claviceps*). Their elastic properties and their technical potential are high: the dragline silk of a spider can stretch by ca. 30 % before breaking and has been explored as a component of bulletproof vests ("BioSteel™"). Like silk, spidroin is formed from repeating polypeptides. It has been expressed in *E. coli, Pichia pastoris* and transgenic animals, using synthetic genes whose codon usage had been optimized for the host organism. Yields of recombinant spidroins, however, are still limited to several g L^{-1} fermentation broth. The greatest challenge for the manufacture of "spider silk" is the preparation of sufficiently long fibers that can be spun in technical devices. An appropriate engineering solution, simulating the spider's spinneret, has already been realized (AMSilk, Spiber).

Adhesion proteins. Mussels, e. g., *Mytilus edilis*, adhere to smooth surfaces such as crab shells by means of a polyamide that hardens under water; by this trick, they have spread over large distances. The M_W of the precursor protein of this natural adhesive is about 130kDa, and the protein is composed mainly of hydrophilic amino acids such as tyrosine, serine, threonine, lysine, and proline. During secretion, the tyrosine and proline residues undergo posttranslational hydroxylation, forming 3- or 4-hydroxy-L-proline and o-hydroxytyrosine (DOPA). Once secreted, these residues, in the presence of oxygen, form chinoid structures, which initiate polymerization of the peptide chains. With an optimized synthetic gene cassette, good expression of the precursor protein was achieved in *E. coli*. The genes for the posttranslational steps involved in adhesive formation, however, cannot be cloned into *E. coli*. As a consequence, the precursor protein is isolated and oxidized by fungal tyrosinase to the polyhydroxylated peptide. L-ascorbic acid is then added to prevent further oxidation. A commercial adhesive based on this material is under development for dental applications.

"Bio-Nylon". Petrochemically-produced nylon is composed of two C-6 building blocks: adipic acid (→152) and hexamethylene diamine (PA 66). Pentamethylene diamine can be produced in a fermentation process by decarboxylation of L-lysine (→128) (BASF, Ajinomoto) and polycondensed with adipic acid to PA 56 whose properties are similar to PA 66. Adipic acid is petrochemically produced from cyclohexene or cyclohexanol, but biotechnological pathways have already been established to produce this dicarboxylic acid from glucose or fatty acids (→152) (Genomatica, Verdezyne). A "green" nylon PA 66 is thus technically feasible, but presently still too expensive. Longer chain dicarboxylic acids such as sebacic acid (C-12) (from ricinoleic acid by alkaline processing) or dodecane dicarboxylic acid (by terminal oxidation of lauric acid, →152) are already commercially available. Thus, PA 510, 512, 610, 612 and other polyamides could in principle be produced from biological building blocks. The higher prices of these materials are, however, not yet accepted by the markets in return for advantages in manufacture or applications.

Silk



Fibroin, secondary structure

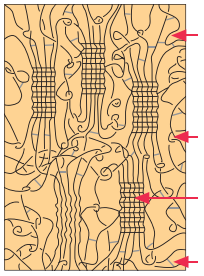
Main component of fibroin

(gly–ser–gly–ala–gly–ala)$_n$-spacer

compressive strength 1×10^9 N m^{-2}
extensibility 15%
tensile strength 350 MPa

Bombyx mori („silkworm") in silk cocoon

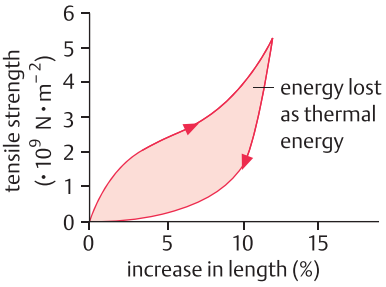
Dragline silk of spiders (*Nephila claviceps*)



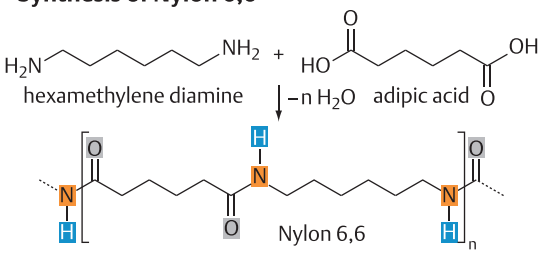
main component of dragline silk

(gly–pro–gly–gly–x)$_{3–63}$–(gly–gly–x)$_{12}$-spacer

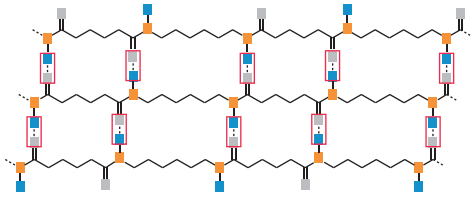
compressive strength $4 \cdot 10^9$ N·m^{-2}
extensibility 500%
tensile strength 2000 MPa



Synthesis of Nylon 6,6



hydrogen bond superstructure in Nylon 6,6



□ hydrogen bonds

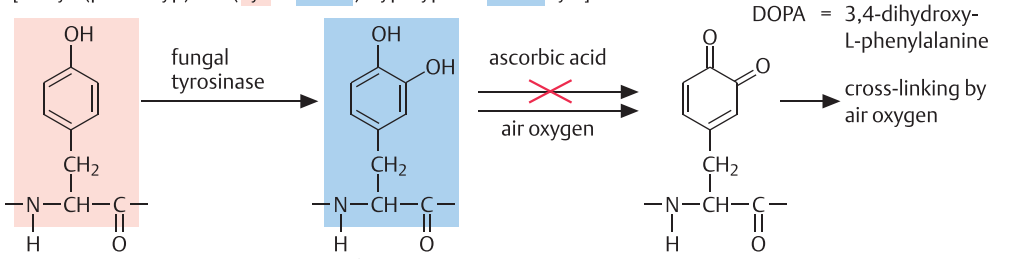
	diamine		diacid	
nylon 6,6	hexamethylene diamine		adipic acid	
nylon 5,6	pentamethylene diamine	decarboxylation of L-lysine (BASF, Ajinomoto)	adipic acid	metabolic pathways from glucose (Genomatica, Verdezyne)
nylon 6,10	hexamethylene diamine		sebacic acid	
nylon 5,12	pentamethylene diamine	decarboxylation of L-lysine (BASF, Ajinomoto)	dodecane dicarboxylic acid	terminal oxidation of dodecanoic acid (Cognis, Cathay, Hilead)

red: petrochemical building blocks, green: biological building blocks

Adhesion protein from common mussel (*Mytilus edulis*)

-[ala-lys-(pro or hyp)-ser-(tyr or DOPA)-hyp-hyp-thr- DOPA -lys-]

hyp = 4-hydroxy proline
DOPA = 3,4-dihydroxy-L-phenylalanine



precursor protein by genetic engineering

activated prepolymer, stabilized by ascorbic acid

Polysaccharides

General. Several polysaccharides such as starch (→176), cellulose (→182), gum arabic, guar, pectins, alginates (→18) and agar have found applications in food manufacture for thickening or stabilization. Xanthan is also used in oil exploration. Although most polysaccharides are isolated from plants or marine algae ("renewable resources"), some are manufactured as extracellular products by microbial fermentation. Because the economics of such processes compare unfavorably to the preparation of plant or marine polysaccharides, their use is limited to specialty areas. The most important microbial polysaccharides are xanthan, dextran and bacterial cellulose. Hyaluronic acid is a component of the extracellular matrix of vertebrates and used in medicine and cosmetics.

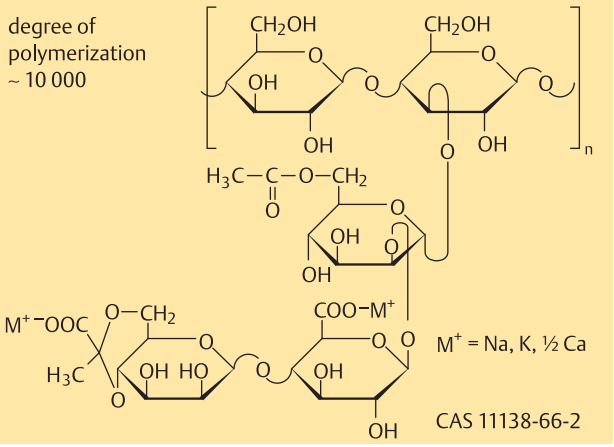
Xanthan is an acid heteropolysaccharide, formed by the prokaryote *Xanthomonas campestris*, a plant pathogen. It is composed of 5 hexose residues as repeating units, and its molecular weight is in the range of $1.5–2 \times 10^6$ Da. The number of pyruvate residues may vary, but this has little influence on the viscosity of the biopolymer. Xanthan viscosity is insensitive to the presence of electrolytes, and the biopolymer exhibits pseudoplastic properties (the viscosity of a xanthan solution decreases reversibly on increasing shear), and thus is very convenient to use in industrial processes requiring gelation. Its most important application is as a thickener in processed food (e. g. salad dressings). The use of xanthan in tertiary oil recovery, for polymer flooding of salt-containing soil formations, was widely studied but, due to the drop in oil prices and improved recovery technology, has not become commercially successful. Xanthan is used in drilling "muds" for oil exploration and development. Fermentation is carried out in a batch mode (C-source: glucose or sucrose; N-source: peptone, ammonium nitrate or urea). The lac operon of *E. coli* has been expressed in *X. campestris*, resulting in strains that can form xanthan from whey, a waste material. This process, however, is not yet commercially competitive. During fermentation, the formation of xanthan is indicated by a steep rise in viscosity to 10,000cP. A special stirrer configuration is crucial if sufficient O_2 transfer in this viscous medium, as required for high yields, is to be achieved. Isolation of the product from the broth is usually done by precipitation with 2-propanol. At present (2013), > 100,000 t y^{-1} of xanthan are produced.

Dextran. Dextrans are used as plasma expanders in medicine and, due to their well defined pore size both in the native form and after chemical derivatization, for the purification of proteins (→106). Dextrans are glucans, built from D-glucose through mostly α-1,6-glycosidic linkages. Their molecular mass reaches 5×10^7 Da. They are produced as extracellular products by several microorganisms, e. g., by *Streptococcus mutans* in the oral cavity of humans, leading to dental plaque. For the technical manufacture of dextran, *Leuconostoc mesenteroides* is usually employed. It produces ca. 200 g L^{-1} dextran in 24h from saccharose, which can be precipitated from the broth by addition of ethanol. If partially hydrolyzed by the addition of acid, dextrans of different molar masses can be isolated by fractionated precipitation. Dextrans of M_W 75kDa are used as plasma expanders; those of M_W 40kDa are used as antithrombolytic agents in surgery.

Other microbial polysaccharides. *Pseudomonas aeruginosa* and *Azotobacter vinelandii* form microbial alginates (→18) whose compositions resemble the algal products. *Gluconoacetobacter xylinus* and others convert sugars and other carbon sources into up to 50 % (w/w) of thin cellulose fibers, associating in microfibrils that can be formed into sheets of high elasticity and purity. These have been used in, e. g., wound dressing and skin transplantation, but also for membranes in high-grade earphones. Some fungi form scleroglucan, a polysaccharide from β-1,3-linked glucose residues containing a lateral β-1,6-linked glucose unit. Similar to gellan from *Sphingomonas elodea*, it exhibits pseudoplastic properties similar to xanthan and is used, in small quantities, as a gelling agent in food products. Pullulan, an α-1,4-linked glucan containing ca. 10 % of α-1,6-glycosidic bonds, is synthesized by various bacteria. It can be processed into films that are impermeable to O_2 and thus has been investigated for the protection of O_2-labile foods or other materials. However, the high production cost of these polymers has so far prevented their use on a large scale.

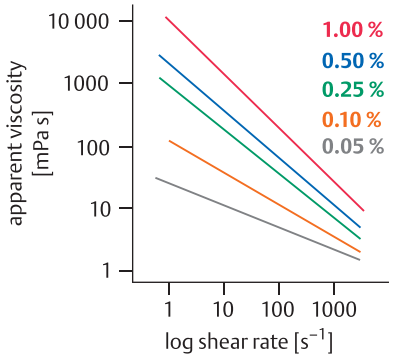
Xanthan

degree of polymerization ~ 10 000



CAS 11138-66-2

pseudoplastic properties



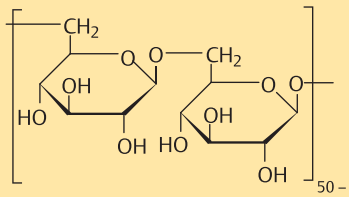
0.05 – 1 % dispersion of xanthan in water, 25 °C

Fermentation and recovery

preculture	bioreactor	recovery	
*Xanthomonas campestris**, glucose or saccharose, N-source	> 120 m^3, 40 – 80 h, 28 °C, pH 7.0, special stirrer equipment due to high viscosity	pasteurization, precipitation with ethanol or 2-propanol	~ 30 $g\ L^{-1}$ in 60 h

**X. campestris* is a plant pathogen and must be cultivated under risk group 2 safety conditions

Dextran



1,6-α-D-glucose, degree of polymerization ~ 28,000
1,2-, 1,3- and 1,4-linkages occur as well

CAS 9004-54-0

biosynthesis

$(1,6\text{-}\alpha\text{-D-glucosyl})_n$ + saccharose

↓ dextran sucrase

$(1,6\text{-}\alpha\text{-D-glucosyl})_{n+1}$ + fructose

Fermentation and recovery

bioreactor	recovery	hydrolysis, fractionation	
Leuconostoc mesenteroides 1) growth 2) production phase: saccharose as C-source, 23 °C	precipitation using ethanol, acetone or methanol	dextrans of varying molecular weight	100 $g\ L^{-1}$ in 24 h yield ~ 45% to saccharose

Market data and applications

polysaccharide	production (t)	price (US$/$kg^{-1}$)	microorganism	applications
xanthan	>100,000	10 – 14	*Xanthomonas campestris*	food additive, oil exploration
dextran, dextran derivatives	2000 600	35 – 390 400 – 2,800	*Leuconostoc mesenteroides*	plasma expander, food additive, biochemicals
bacterial cellulose	small	~ 100	*Gluconacetobacter xylinium*	wound dressing, lead/phones

other microbial polysaccharides with some economic importance:
alginate (*Azotobacter*), curdlan (*Agrobacterium*, *Rhizobium*), gellan (*Sphingomonas*), pullulan (*Pullularia*), β-glucans (fungal origin)

Biosurfactants

General. Some microorganisms, if exposed to alkanes, plant oils, or even sugar substrates, form surface-active agents that are often termed "biosurfactants". Compared with synthetic surfactants that are produced in amounts of several million tons, biosurfactants are much more expensive and thus are used only in limited specialty areas, e. g., in cosmetics. Due to their high biodegradability, they are being studied as an environmentally "friendly" means of decontaminating oil-polluted water, soil, and tidal shallows.

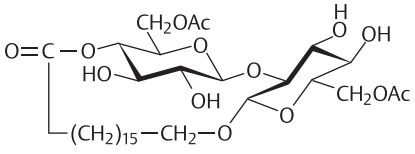
Biosurfactants are formed by several bacteria and yeasts when grown on alkanes or plant oils. Well-studied biosurfactants include the sugar esters rhamnose-, trehalose- and mannosyl erythritol-lipid, the lipopeptide surfactin, and the heteropolysaccharide emulsan. Sophorose lipids are well-known examples of yeast biosurfactants. They are formed by *Starmerella bombicola* in yields of >400 g L^{-1} if the organism is grown on triglycerides. The gene cluster for the formation of this product was cloned and sequenced. It contains five genes: a P450 monooxygenase for the formation of the hydroxy fatty acid, two glucosyl transferases, an acetyl transferase and a transporter protein. Using metabolic engineering techniques, the yield of this product was improved. The mixture of hydroxy acid glycolipids and their lactones can be separated from the fermentation broth by flotation or solvent extraction. The purified products tend to form micelles, and their CMC (critical micelle concentration, a value indicating surfactant activity) is in a range typical of synthetic nonionic surfactants. The fungus *Pseudozyma hubeiensis* SY62, in the presence of triglycerides, also forms large quantities of acetylated mannosyl erythritol lipids (MEL-B). Its 18.4 Mb genome was recently sequenced. Other organisms studied are the fungus *Ustilago maydis* and *Cryptococcus humicola* that produces cellobiose lipids. Rhamnolipids are synthesized by some Pseudomonas strains and can be prepared in yields of ca. 100 g L^{-1}. They have been tested as biosurfactants in microbial enhanced oil recovery (MEOR) (producing the sugar lipid in-place at the drilling hole) and for cleaning up oil pollution from tidal shallows. Good results were obtained from a technical but not an economical point of view. A fermentation broth from a non-pathogenic strain of *Pseudomonas aeruginosa* that contained 120 mg/L rhamnose lipids absorbed 27 % of the crude oil on sand in a laboratory experiment. This strain synthesized rhamnose lipids with sunflower oil as the only C-source under N-limiting culture conditions. *Rhodococcus erythropolis* has been shown to secrete trehalose tetraesters. Depending on the type of hydrophobic substrate, the chain length of these glycolipids can be modulated. Emulsan is a heteropolysaccharide (lipopolysaccharide) synthesized by *Acinetobacter calcoaceticus* in the presence of triglycerides. It can be produced by fermentation and isolated from the culture broth by solvent extraction. If added to a suspension of oil in water, it acts as an oil-in-water emulsifier, thus significantly reducing the viscosity of the oil. It has been used in small quantities for improving the flow of oil in pipelines and for cleaning oil trucks and barges. *Bacillus subtilis*, if grown on hydrophobic substrates, forms the acylated heptapeptide surfactin that exhibits a high CMC, but also is hematotoxic to mammals and aquatic organisms. Yields in an optimized fermentation procedure may be as high as 110 mg L^{-1} but are orders of magnitude lower than yields of yeast biosurfactants such as sophorolipid.

Industrial Products. Presently (2014) two biosurfactants are industrially produced: sophorolipid from *Starmerella bombicola* by Ecover (Belgium), and MEL-B from *Pseudozyma hubeiensis* by Toyobo in Japan. Ecover uses sophorolipids to formulate a range of household detergents and features the excellent environmental properties and ecobalance of this nonionic detergent. However, this product competes with APG, alkylated polyglucosides of similar structure, which are manufactured from fatty alcohols and glucose by chemical catalysis. Their ecobalance (→330) is equally good, but the product is cheaper to produce. MEL-B, an acetylated mannosyl erythritol lipid, is manufactured by Toyobo (Japan) and used by Kanebo, a cosmetic company, as a natural ingredient in skin care products. It has similar properties to ceramide, a natural wax. The strong moisturizing effect of MEL-B was shown in appropriate skin models. Variants of MEL-B have been reported to form liquid crystals.

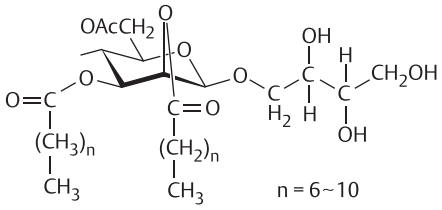
Biosurfactants

	organism	structural elements
sophorose lipid	*Starmerella bombicola*	sophorose, hydroxy fatty acids
cellobiose lipids	*Ustilago maydis*	cellobiose, fatty acids
mannosyl-erythritol lipids	*Pseudomonas hubeiensis*	mannose, erythritol, fatty acids
rhamnose lipids	*Pseudomonas aeruginosa*	rhamnose, β-hydroxydecanoic acid
trehalose lipids	Corynebacteria, Arthrobacter	trehalose, long-chain wax esters
corynomycolates	Corynebacteria, Arthrobacter	mycolic acid esters of mono-, di-, and trisaccharides
emulsan	*Acinetobacter calcoaceticus*	polyanionic heteropolysaccharide, $M_R \sim 10^6$ Da
surfactin	*Bacillus subtilis*	acylated heptapetide

sophorose lipid

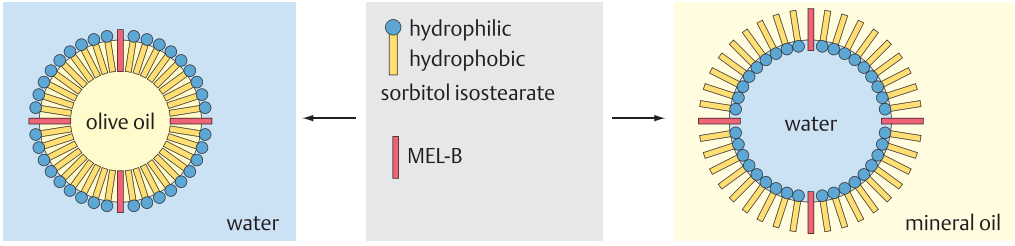


mannosyl erythrit lipid (MEL-B)



organism	fermentation	recovery	
Starmerella bombicola	plant oil, gutter oil	centrifugation, solvent extraction	up to 400 g L^{-1} broth

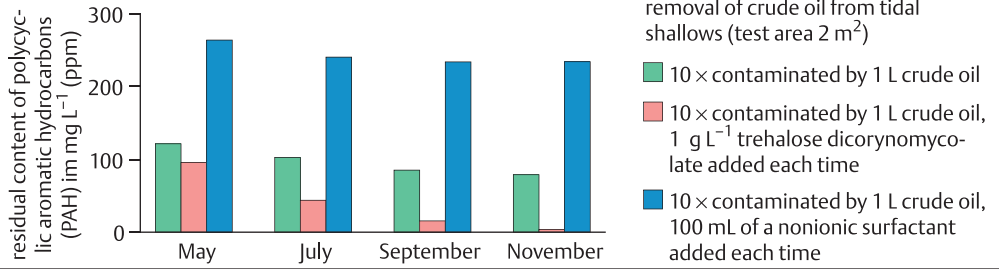
Moisturizing effect of biosurfactant MEL B



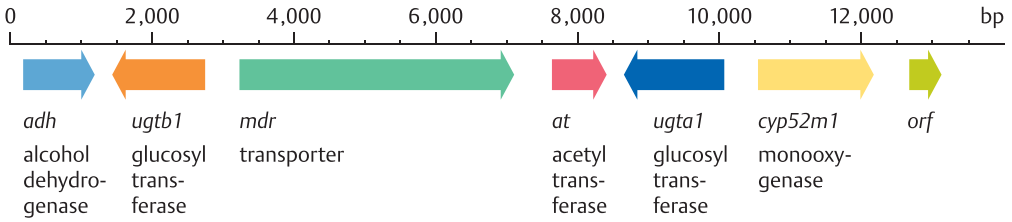
oil-in-water (o/w) emulsion

water-in-oil (w/o) emulsion

Biosurfactants and marine pollution



Gene cluster for sophorose lipid biosynthesis in *S. bombicola*



Fatty acids and their esters

Fats and oils are produced from animals and oil plants in annual quantities of 190 million t (2012). Their dominant use is in food and feed products, but a small part is used in chemical industry for the synthesis of surfactants, emulsifiers etc. ("oleochemistry"). Animals and plants use phosphoglycerides as membrane components and triglycerides for energy storage. Most fatty acids in triglycerides are built from chain length of preferentially C-12 to C-20, either saturated or with 1 – 3 double bonds. Fish oils contain longer and higher unsaturated fatty acids such as docosahexaenoic acid (DHA) and eicosapentaenoic acid (EPA), often used as food additives (→162). Through biotechnology, triglycerides can be hydrolyzed, re-esterified (transesterified) or synthesized from glycerol and fatty acids using lipases. Some yeasts, fungi and algae store up to 80 % of cell moist mass as triglycerides that can be used directly, or after modification, as energy chemicals ("biodiesel") or food additives.

"Biodiesel" is a mixture of fatty acid methyl or ethyl esters, which are produced through transesterification of triglycerides with methanol or ethanol. They can be blended with standard diesel fuel. NaOH/methanol is used for transesterification, but lipases such as the enzyme from *Yarrowia lipolytica* have also been used. Enzymatic transesterification with methanol is advantageous when waste cooking oil ("gutter oil") is used for biodiesel synthesis. Biodiesel is also under discussion as a kerosene substitute in jet fuel, and several test flights have been made.

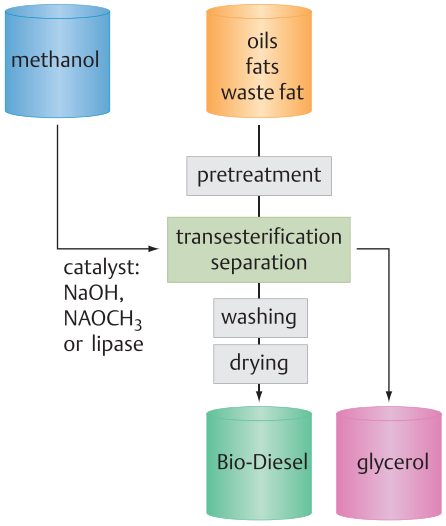
Cocoa butter is obtained from cacao beans, the fruit of the tropical cacao tree. For manufacturing, roasted cacao beans are grounded, and after gentle warming nearly equal amounts of cacao butter and cacao paste are obtained. Cacao paste is the base material for cacao powder. Cacao butter is composed from a mixture of triglycerides with palmitic, stearic and oleic acid as the predominant fatty acids. Its melting point is ~37 °C, the human body temperature. As a consequence, cacao butter melts quickly on the tongue, as required for good chocolates. It is also used in suppositories for the rectal application of drugs. A cocoa butter substitute can be enzymatically prepared through transesterification of palm oil (main fatty acids: palmitic and oleic acid). According to food legislation, it is not considered a natural compound and must be labeled.

Value-added glycerides through enzyme catalysis. Regio-specific lipases such as the enzyme from *Rhizomucor miehei* (1,3-specific) can be used to prepare a diet food oil (ECONA®) composed largely of 1,3-digly-cerides which are not resorbed in the intestine and thus do not lead to fat storage. The process starts from glycerol and free fatty acids. BETAPOL® is a triglyceride mainly composed of glycerol 1,3-oleate 2-palmitate ("OPO"), resembling the fat of breast milk; it is used in milk powder formula for babies which are not breast-fed. The preparation follows a lipase-catalyzed transesterification of triglycerides of suitable composition or through lipase-catalyzed addition of oleic acid (or transesterification of oleic acid ethyl ester) to 2-palmitoyl monoglyceride. Another example for a synthetic triglyceride is TONALIN®, an oil that is advertised as a sports diet and slimming agent as it inhibits the activity of lipoprotein lipase and thus prevents fat storage. It contains CLA, a mixture of *cis*-9-*trans*-11- and *trans*-10-*cis*-12-linolenic acid.

Oil-accumulating microorganisms have been explored for many decades, e. g., for the production of food and feed oils in war times. Yeasts such as *Mortierella alpina*, *Rhodosporidium toruloides, Lipomyces starkey* and algae such as *Botryococcus braunii* or *Neochloris oleoabundans* can store up to 80 % of their moist mass as triglycerides (composed of mostly higher unsaturated fatty acids) and thus be exploited for the production of food additives or biodiesel. The genomes of some of these organisms were already sequenced, and there is ongoing research in metabolic engineering and other modifications of such strains.

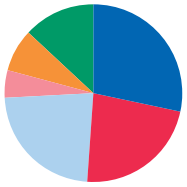
Highly unsaturated fatty acids with an (ω-3)- or (ω-6)-double bond such as DHA (docosahexaenoic acid, 22:6(ω-3)), EPA (eicosapentaenoic acid, 20:5(ω-3)) or arachidonic acid 20:4(ω-6) occur in fish oil or peanut oil and are used as a health food additive against atherosclerosis. Starting from standard fatty acids such as oleic acid, biosynthetic steps include chain elongation and the introduction of double bonds. Most of these enzymes have been cloned and functionally expressed in yeasts such as *Mortierella alpina*, leading to recombinant organisms that produce high amounts of EPA, DHA or arachidonic acid.

Bio-Diesel



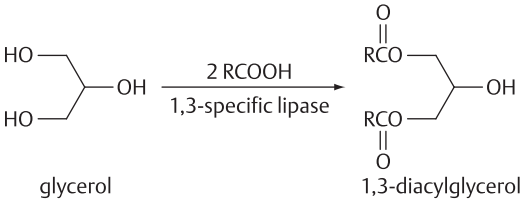
Global production of oils and fats

2011/2012: 182.9 million tons



palm oil	28.2%
soy oil	22.8%
others and animal fats	23.2%
palm kernel oil and coconut oil	4.9%
sunflower oil	8.0%
rapeseed oil	12.9%

Enzymatic synthesis of a 1,3 diglyceride food oil



glycerol

1,3-diacylglycerol

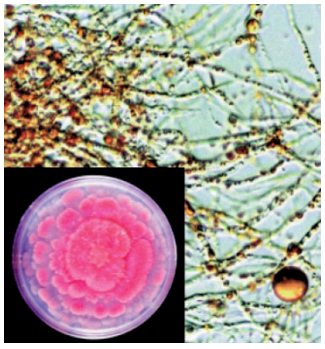
R = oleic, linoleic and others

Melting point of different triglycerides

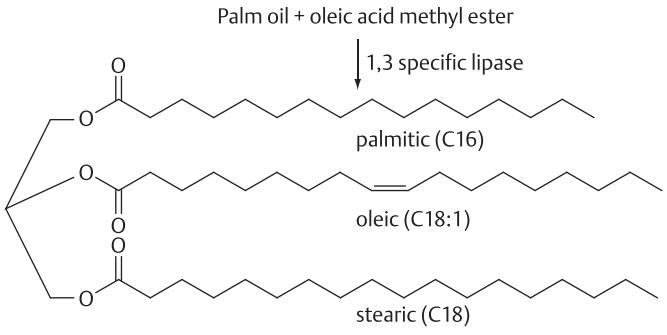
"tallow" (3-stearyl-1,2-dipalmitin)	62 °C
"cocoa butter" (1-palmityl-2-oleyl-3-stearin)	37 °C
"palm oil" (1-stearyl-2-palmityl-3-olein)	39 °C
triolein	5.5 °C

Cocoa beans and cocoa butter

Mortierella alpina



Enzymatic synthesis of cocoabutter



Fats through enzymatic transesterification

fat	raw materials	enzyme
1,3-diglyceride rich oil (ECONA)	glycerol, fatty acids	1,3-specific lipase
2-palmitoyl-1,3-oleyl-glycerol (BETAPOL)	appropriate triglycerides	1,3-specific lipase
2-γ-linolenic-1,3-caproyl-glycerol (TONALIN)	2-γ-linolenoyl-monoglyceride, caproic acid	1,3-specific lipase
1-stearoyl-2-palmitoyl-3-oleyl-glycerol (cocoabutter substitute)	palm oil, stearic acid	lipase

Biotransformation

General. Biotransformations are key functions of organisms and serve for the biosynthesis or biodegradation of metabolites (anabolism and catabolism) or for the detoxification of toxic or unnatural (xenobiotic) compounds. Each step in metabolism is catalyzed by an enzyme. In biotechnology, biotransformation is usually a synonym for biocatalysis, or the transformation of natural or synthetic precursors (educts) into products of higher value. Technical biotransformations are carried out with microorganisms, mammalian cells, or plant cells in a bioreactor (fermentation) or with isolated enzymes or cells (that can be immobilized on carrier materials). These advances in genetic engineering have revolutionized biotransformations. Today, not only can genes coding for metabolic steps in animals or plants be transferred into microorganisms, allowing for simpler access and use, but metabolic pathways not occurring in nature that still lead to high yields of valuable products can be artificially assembled in microbes ("synthetic biology"). Whether a given biocatalytic process is termed "biocatalysis", "biotransformation", "fermentation", or "enzyme catalysis" often depends on personal preference, since a single or several enzymes may be required in all these processes. The use of isolated enzymes can simplify a process, as temperature tolerance is usually better, sterile conditions are not required, and diffusion of educt and product are unhindered. However, using an isolated enzyme may not be an option if it is too expensive to isolate, is unstable, or requires auxiliary enzymes and cofactors.

Microorganisms are used for producing natural metabolites (e. g., glutamic acid) (→126) and also for the biotransformation of unnatural substrates (e. g., in the 11β-hydroxylation of synthetic steroid derivatives) (→252). Being enzyme-catalyzed reactions, these transformations generally proceed in a regio- and stereoselective manner. Since genes or whole gene cassettes from other organisms can be cloned and expressed in a host microorganism, the possibilities for biotransformations have expanded dramatically (e. g., for the microbial production of indigo). Metabolic engineering (→318) and protein engineering (→198), as well as the discovery of novel enzymes and pathways (→320) through genome sequencing, have further helped to expand the industrial use of biotransformation reactions.

Mammalian cells are used industrially in the manufacture of pharmaceutical proteins (cell fermenters), but are too expensive for single-step biotransformations. They are being investigated in medicine, e. g., for use as an "artificial liver" that transforms and eliminates toxic metabolites found in blood.

Plant cells have been studied for biotransformation, e. g., for the 12-hydroxylation of β-methyl digitoxin to β-methyl digoxin with cell cultures of *Digitalis lanata* (→278). Compared to chemical synthesis or recombinant microorganisms, plant cell cultures have had only limited success in industrial processes, e. g., for the synthesis of taxol (Paclitaxel™) with cell cultures of *Taxis brevifolia* (→278).

Enzyme catalysis. Usually, enzymes are used in single-step biotransformation processes. In most processes, isolated enzymes in free or immobilized form (→102) are used, but if the enzyme proves too expensive to isolate, an active enzyme in a whole inactivated microorganism may be used (e. g., glucose isomerase in Streptomyces cells) (→180). Most industrial examples of enzyme biotransformations involve hydrolases, since they do not require a cofactor and often catalyze regio- and stereoselective reactions (see Applied Enzyme Catalysis). Enzymatic isomerization reactions, as well as addition reactions to double bonds, carbonyl groups, or activated CH bonds have also found industrial applications.

Recombinant metabolic pathways. This technology has advanced rapidly over the last decade (see Metabolic Engineering, Synthetic Biology, →320). An early example is the production of indigo for which naphthalene dioxygenase from *Pseudomonas* sp. was cloned into *E. coli*, followed by optimization of the novel pathway through metabolic engineering of the host organism. Genes coding for the transformation of glycerol to propane 1,3-diol in *Klebsiella pneumoniae* were engineered into *E. coli*, allowing for the commercial production of this polyester building block (→142). Many other reactions of this type are described in this book, e. g., the preparation of D- or L-amino acids at will from synthetic hydantoins using a combination of stereospecific hydantoinases and carbamoylases, co-expressed in an *E. coli* host strain optimized for secretion (→132).

Biotransformations

biotransformation/biocatalysis
synthesis or transformation of a compound

- **fermentation** living cells in a bioreactor
- **resting or immobilized cells** cells in suspended or carrier-bound form
- **enzyme catalysis** one or several isolated enzymes

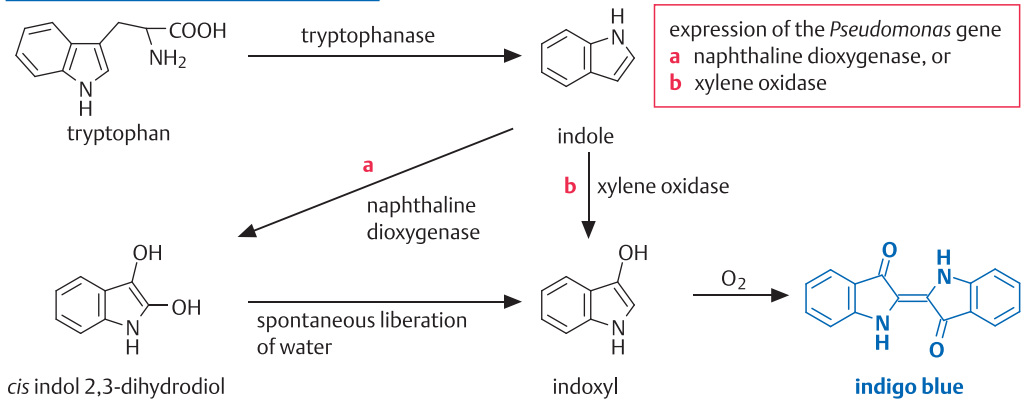
Biotransformation reactions (examples)

	organism/enzyme	type	company
D-sorbitol → L-sorbose	*Acetobacter suboxydans*	F	DSM
phenyl-D-lactate → 4-hydroxy-phenyl-D-lactate	*Beauveria gossypii*	F	BASF
fumaric acid → L-aspartic acid	*Escherichia coli*	IC	Ajinomoto
D-glucose → D-fructose	glucose isomerase from *Streptomyces* sp.	IC, IE	ADM, Tate & Lyle
D,L-acetoxymethoxyphenylethylamine → L-phenylethylamine	lipase from *Burkholderia cepacia*	IE	BASF
tryptophan → indigo	recombinant strain of *E. coli*-Stamm	RIC	Genencor
glucose → propane 1,3-diol	*E. coli* with *Klebsiella* pathway	RIC	DuPont, Genencor

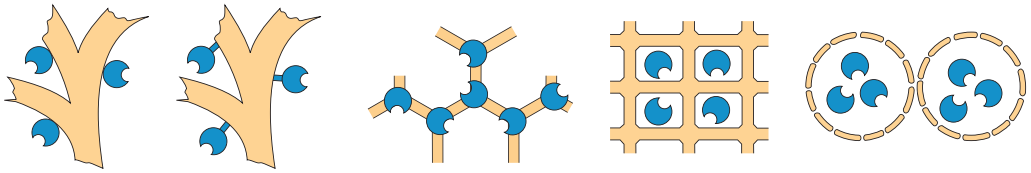
F: fermentation, IC: immobilized cells, IE: immobilized enzyme, RIC: recombinant immobilized cells

Indigo from recombinant *E. coli*

genetically deregulated *E. coli* which overexpresses tryptophane and indole



Immobilization of enzymes and cells



adsorbed to carriers | covalently linked to carriers | cross-linked | encapsulated in polymers | microencapsulated

Technical enzymes

General. Starting about 100 years ago, various enzymes (→102) isolated from animals, plants, and microorganisms have been developed into important components used in technical processes and into analytical reagents. Around 1970, immobilized enzymes were developed for industrial biocatalysis (enzyme transformation) (→164). Genetic engineering has further increased this potential by providing purer or genetically modified enzymes, e. g., by protein engineering (→198) or directed evolution to adapt enzyme catalysis to technical needs.

Enzyme classification. Following international agreements, enzymes are divided into six classes, based on their function (→30). Several thousand enzymes of various functions are known, and for most of them additional variants have been isolated from different organisms. Usually their properties are unfavorable for use in industry. For example, about one third of all enzymes function in the environment of biological membranes and are rather unstable in isolated form. The activity of most oxidoreductases, transferases, ligases, and synthases requires cofactors such as NADH, ATP, or coenzyme A, which are expensive and must be regenerated for economic reasons. Hydrolases and isomerases do not share these disadvantages and thus are preferred enzymes in industrial applications. For analytical and diagnostic applications, however, the high selectivity of enzymes justifies a high prize, leading to the use of all kinds of enzymes.

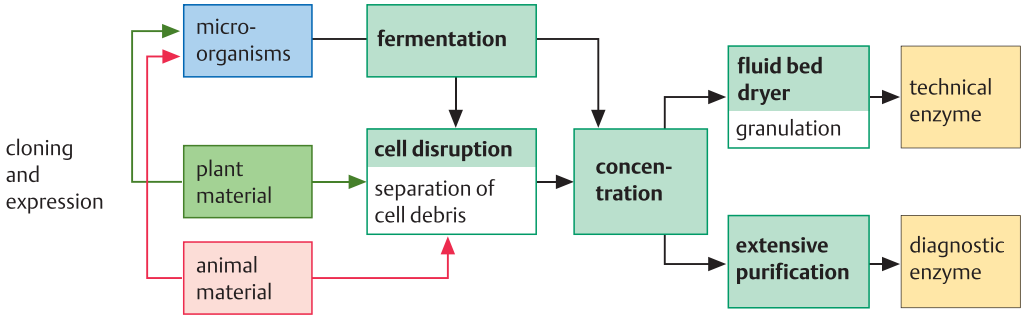
Manufacture. Production methods vary greatly, depending on the origin of the enzyme (animals, plants, microorganisms), the intended use (degree of purity required), and production scale. Other factors that largely determine the individual isolation and purification protocols are the properties of the desired enzyme (soluble or membrane-bound, stable or labile). If large quantities of an enzyme are required for a technical application (e. g., microbial proteases for use in detergents), extracellular enzymes produced by fermentation are preferred, as are simple steps for downstream processing (→104) such as separation of cells, concentration of fermentation broth by precipitation or ultrafiltration, gentle drying in a spray or vortex dryer, or adding stabilizing compounds for finishing. The enzyme preparations obtained by such processes are usually of low purity and may be contaminated with other enzymes (which actually may be useful for the desired application, e. g., in food processing). Enzyme preparations that are used in therapy (e. g., tPA (→230), DNase (→240)) or for diagnostic purposes (→0254 must be highly pure. They are often produced inside cells and, because they achieve a higher price in the market, are usually isolated and purified in several chromatographic purification steps (→106) that eventually lead to a single enzyme activity. The purity of the enzyme preparation is monitored through 1) determining the specific activity at each purification step, 2) determining the decrease in undesired side-activities, and 3) electrophoretic methods. Genetic engineering has revolutionized enzyme production, and many enzymes used in technology today are produced by fermentation procedures based on recombinant host microorganisms. A particular advantage of these processes is that fewer side-activities are present, and a pure enzyme preparation can be obtained in fewer chromatographic steps, producing a reduced amount of waste materials.

Registration. (→334) Food additives are considered "natural" and need not be labelled if they are produced from natural materials through enzymes which do not remain in the final product. An example is isoglucose. Though enzymes are also natural products, their applications in food and human therapy require registration. In Europe, registration of food and feed enzymes may proceed after dossiers have been submitted that contain detailed information about the manufacturing process, efficiency, and health- and safety-related issues. The agency in charge is the European Food Safety Agency (EFSA), but in countries such as France or Denmark, registration is still on the national level, depending on the use of the enzyme as a food additive or a processing aid. Recombinant enzymes are subject to special regulations. Technical enzymes, e. g. detergent enzymes, must be registered in Europe by the European Chemical Agency under the REACH program. In the USA, food enzymes that have obtained GRAS status (**g**enerally **r**ecognized **a**s **s**afe) are outlined in a positive list and can be freely used after notification to the FDA. The GRAS status requires that the organism has a long history in the production of fermented foods, as recognized by experts.

Enzyme classification

EC-number	name	coenzymes	example
1.x.y.z	**oxidoreductases**		
1.1.y.z	react with CH—OH	NAD^+, $NADP^+$, PQQ	alcohol dehydrogenases
1.1.3.z	react with CH—OH	FAD^+	glucose oxidase
1.3.y.z	react with C—H	heme, Fe^{2+}	steroid 11β-hydroxylase
2.x.y.z	**transferases**		
2.4.y.z	transfer glycosyl groups		glucosyl transferase
2.6.1.z	transfer NH_3 to carbonyl groups	pyridoxal phosphate	transaminases
3.x.y.z	**hydrolases**		
3.1.y.z	hydrolyze ester bonds		lipases, esterases
3.2.y.z	hydrolyze glycoside bonds		α-amylase
3.4.y.z	hydrolyze peptide bonds		subtilisin, trypsin
4.x.y.z	**lyases**		
	catalyze addition or elimination reactions		citrat lyase
4.3.y.z	add or eliminate NH_3 to/ from C=C-double bonds		aspartase
5.x.y.z	**isomerases**		
5.1.y.z	racemize D- and L-amino acids		alanine racemase
5.3.y.z	intramolecular oxidoreductases		xylose (glucose) isomerase
6.x.y.z	**ligases**		
6.2.y.z	C-S coupling	ATP, CoASH	acetyl-CoA synthetase

Preparation of enzymes



Registration

origin	examples	recommendations
animal organs	**pancreas extracts, rennin, pepsin**	**good manufacturing practice**
plant material	**papain, bromelain**	**good manufacturing practice**
microorganisms		
a) traditionally used for the production of food products	*Bacillus subtilis*, *Aspergillus niger*, *A. oryzae*, *Mucor javanicus*, *Rhizopus* sp., *Saccharomyces cerevisiae*, *Kluveromyces fragilis* and *K. lactis*, *Leuconostoc oenus*, organisms from starter cultures	**USA**: GRAS positive list for type a) microorganisms (GRAS = generally recognized as safe) **EU**: for food and feed enzymes, approval by the European Food Safety Agency, for technical enzymes approval under REACH by the European Chemical Agency
b) enzymes from well known microorganisms	*Bacillus stearothermophilus*, *B. lichenformis*, *B. coagulans*, *B. megaterium*, *B. circulans*, *Klebsiella aerogenes*	

Applied enzyme catalysis

General. Enzymes are used in chemical syntheses due to their favorable energetics and their regio- and stereoselectivity ("green technology") (→170). Enzyme catalysis is performed with isolated enzymes or with optimized enzymes in recombinant microorganisms (see "Biotransformation", →174). Those enzymes in isolated reactions that do not require coenzymes are preferred. The most important examples are hydrolases, but some isomerases and oxidoreductases with firmly bound coenzymes are also used. Today, many enzymes are available in recombinant form. They can be prepared in large quantities, free of contaminating enzymes, and they can be optimized for industrial or diagnostic use by methods such as protein engineering (→198). The discussion here uses the numbering of enzyme nomenclature (→166).

Oxidoreductases. By 2014, about 2,000 different types of oxidoreductases have been classified. Oxidases usually contain firmly bound FAD as a cofactor and thus can be used for analytical test strips. Dehydrogenases are very valuable for both analytical and synthetic applications, since their reaction leads to equilibria; consequently, they can be applied either for the oxidation of hydroxyl or the reduction of carbonyl groups. For technical applications, however, their expensive coenzymes, usually $NAD(P)^+$ or NAD(P)H, must be added, or an inexpensive coupled reaction for regenerating these coenzymes must be found. A breakthrough occurred when formate dehydrogenase from *Candida boidinii* and NAD(P)H bound to PEG as a coenzyme were used with dehydrogenases in an enzyme membrane reactor (→132), shifting the equilibrium to complete oxidation of the substrate due to formation of CO_2 as the coupled product. In the meantime, the manufacture of recombinant enzymes has greatly advanced and permits the use of coupled enzymes for cofactor regeneration (e. g., den Einsatz gekoppelter Enzymsysteme (z. B. glucose- or alcohol dehydrogenases) (→170). Recently, interest in the technical use of peroxidases, dioxygenases, and P450 monooxygenases (→324) has been growing; they all can oxidize non-activated carbon–hydrogen bonds in a regio- and stereoselective manner. They often contain firmly bound Fe-S clusters or heme cofactors.

Transferases include ~1,700 enzyme types that are not currently used in technical applications.

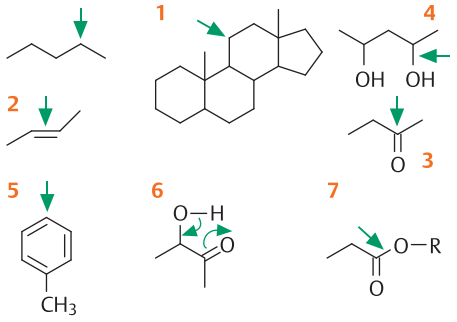
Hydrolases are the most important group of technical enzymes. They comprise ~1,700 types of enzymes; lipases, esterases, amylases, and proteases have found the greatest technical application. If water content and water activity of the reaction mixture is thoroughly controlled, they can be used for regio- and stereoselective formation of ester and amide bonds. One example is *Bacillus stearothermophilus* thermolysin that is used for regiospecific esterification of L-aspartate and L-phenylalanine methylester, providing aspartame sweetener (→130); another example is *E. coli* penicillin amidase that is widely used for synthesis of semisynthetic penicillins from 6-aminopenicillanic acid. Lipase from *Burkholderia cepacia* is used for industrial synthesis of chiral amines from racemic amide precursors (→170), and lipase from *Rhizomucor miehei* is extensively used for manufacturing cocoa butter substitutes (→162). Amino acylase is used for enantioselective hydrolysis of racemic N-acyl amino acids (→132).

Lyases include ~680 examples, and most are cofactor-independent. Aspartase from *E. coli* is used for manufacturing L-aspartic acid from fumaric acid on a technical scale, using immobilized cells instead of the free enzyme (→130). Acrylonitrile hydratase from *Pseudomonas chloraphis* catalyzes the addition of water to acrylonitrile, forming acrylamide, the monomer of polyacrylamides. Oxynitrilases enable addition of HCN in a stereoselective manner to carbonyl compounds, resulting in chiral D- or L-hydroxy acids after hydrolysis of the nitrile group (→132). Using aldolases, e. g., from rabbit liver, various sugars have been synthesized from suitable building blocks. Pectate and pectin lyase are used in the food industry (→186).

Isomerases are a small group of enzymes, comprising just ~280 members. They do not need a coenzyme. Glucose isomerase is used on a large industrial scale for partial isomerization of D-glucose to D-fructose (→180), resulting in isoglucose syrups of high sweetness. The intracellular enzyme is usually used within inactivated, immobilized cells of its producing microorganism, often Streptomyces (→102).

Ligases. All ~200 ligases are ATP-dependent. Presently, they are not used in a technical process with regeneration of ATP.

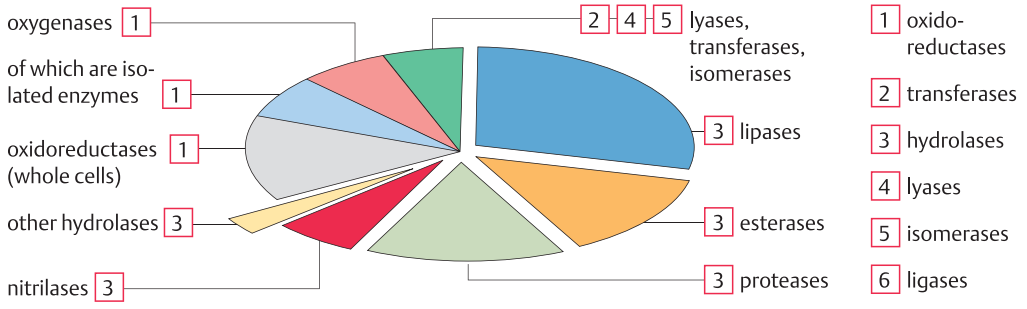
Enzymes vs. chemical synthesis



enzymes are usually superior for

1 regio- and stereospecific oxidations at nonactivated C–H bonds
2 stereoselective addition reactions at activated carbon atoms
3 stereoselective reduction of carbonyl groups
4 regioselective oxidation of hydroxy groups
5 position-specific substitution at the aromatic ring
6 position-specific isomerization
7 regio- or stereoselective hydrolysis or formation of esters under mild conditions

Types of enzymes used in biocatalysis



Enzyme processes in industry (2014, examples)

	process	enzyme	scale (t/year)	company (examples)
3	hydrolysis of penicillin G to 6-amino penicillanic acid	penicillin amidase* from *E. coli*	50 000	DSM Sinochem, Sandoz
3	hydrolysis of *N*-acyl-DL-amino acids to L-amino acids	acylase* from *Aspergillus* sp.	5 000	Evonik, Ajinomoto
4	addition of NH_3 to fumaric acid yielding L-aspartic acid	aspartase* within *E. coli*	10 000	Ajinomoto
3	hydrolysis of starch to D-maltose and D-glucose	α-amylase, glucoamylase	>100 000	several companies
5	isomerization of D-glucose to D-fructose	glucose isomerase* within *Streptomyces* sp.	8 000 000	ADM, Tate & Lyle
3	synthesis of acrylamide from acrylonitrile	nitrile hydratase* within *Pseudomonas chloraphis*	300 000	Nitto Chemicals, BASF
3	hydrolysis of racemic secondary amides to chiral amines	lipase* from *Burkholderia cepacia*	10 000	BASF
3	transesterification of palm oil with stearic acid methyl esters, yielding cocoa butter substitute	lipase* from *Rhizomucor miehei*	1 000	Unilever, Fuji Oil
3	dehalogenation of 1-chloro propane diol	dehalogenase* from thermophilic organisms	under development	Dow Chemicals

*in immobilized form

Regio- and enantioselective enzymatic synthesis

General. As shown for the example of amino acids, enzymatic reactions can often be used to synthesize chiral molecules. In the framework of this short text, only a few important examples can be provided, such as 1) the synthesis of chiral amines using lipases or transaminases, 2) the deracemization of hydroxyl groups using lipases, esterases, ketoreductases, epoxide hydrolyses, dehalogenases or aldolases, and 3) the synthesis of chiral carboxylic acids using esterases, lipases or nitrilases. The partial hydrolysis of nitriles to acid amides will also be described. Many of these reactions are already used in industry, for both the synthesis of end products or building blocks (synthones).

Chiral amines. The enzymatic acylation of racemic secondary amines with methoxy acetic acid proceeds, in the presence of lipases such as the enzyme from *Burholderia cepacia*, with high yield and enantioselectivity, leaving behind a chiral secondary amine of high optical purity. The chiral amide can be hydrolyzed to the undesired chiral amine, racemized and re-entered into the reaction cycle. Through this reaction, a wide range of chiral secondary amines has become industrially available as synthones for pharmaceuticals and agrochemicals. In an alternative reaction, transaminases transform keto compounds into chiral amines by reductive amination with high regio- and enantioselectivity. This reaction requires pyridoxal phosphate as a cofactor. If isopropylamine is used as amino group donor, acetone will be the coupled product of this reaction. The commercial synthesis of Sitagliptin, an antidiabetic drug, is based on this reaction.

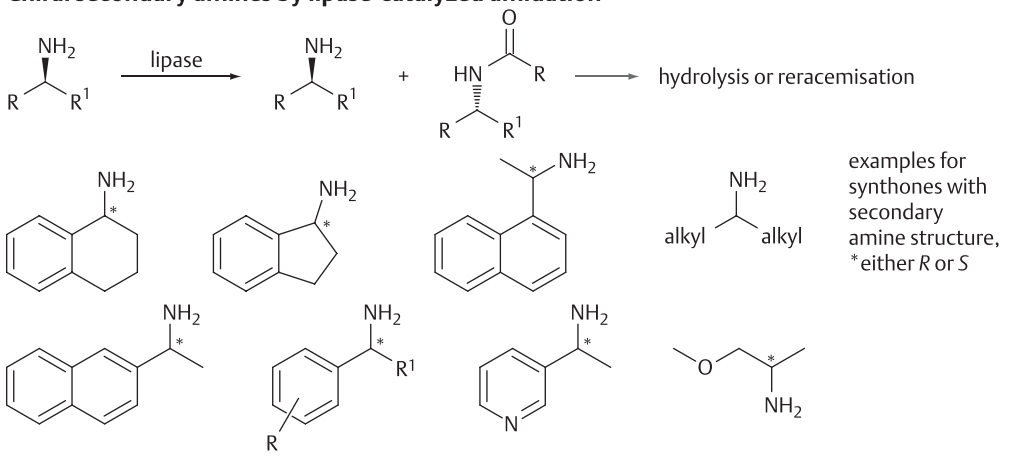
Chiral hydroxy groups. Lipases and esterases split racemic esters usually with high enantioselectivity. For this type of reaction, a wide range of recombinant lipases and esterases have become available, and further optimization for the desired reaction by techniques such as directed evolution or protein engineering (→198) have been described in many cases. Another broadly applicable reaction uses the enantioselectivity of dehydrogenases or ketoreductases. These NADPH-dependent enzymes are mostly applied as a whole-cell catalyst, usually with *E. coli* as the host organism, as this allows the regeneration of NADPH consumed in this reaction by a second recombinant enzyme, e. g., a glucose dehydrogenase or alcohol dehydrogenase, which has been cloned into the host organism. These reactions can also be used for the preparation of chiral hydroxy acids. For the synthesis of some chiral diols, epoxide hydrolases are the enzymes of choice, as they hydrolyze racemic epoxides enantioselectively. Aldolasen may also be used for this type of reaction, as many types of aldolases with various educt- and substrate specificity are available. The skillful synthesis of a chiral vicinal diol by the combined use of a ketoreductase and a haloalkane dehalogenase is commercially used in the synthesis of Atorvastatin (Lipitor®), a drug that lowers cholesterol levels in blood.

Chiral hydroxycarboxylic acids. If an ester molecule contains a prochiral group in its acid part, the enantiomeric product can usually be obtained by lipase or esterase catalysis. Nitrilases are also often used for the synthesis of chiral hydroxy acids, as they hydrolyze only one of the two enantiomers from a chemically produced racemic hydroxynitril. The synthesis of Clopidogrel, a hemorrhage inhibitor, is done by this reaction.

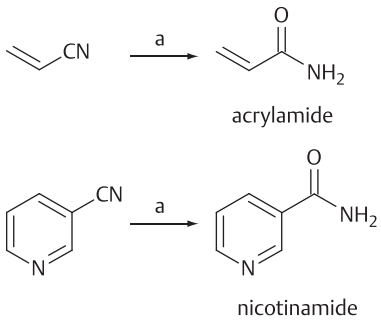
Acid amides. Nitrilhydratases transform nitrils into acid amides. No chiral center is formed in this reaction, but the reaction is of high economic relevance: acrylamide, the building block of polyacrylamide, hitherto produced by chemical catalysis in volumes of several hundred thousand tons, can now be prepared by this reaction in high space-time yields. Using nitrilhydratase from *Rhodococcus chloraphae* in a whole-cell catalyst, hydration of acrylonitril to acrylamide proceeds selectively and under environmentally benign conditions, providing a product of excellent purity and ready for polymerization.

Reduction of alkenes. The asymmetric reduction of C=C bonds is among the very important reactions in organic synthesis, as up to two stereo centers can be formed by this reaction. If enzyme catalysts are to be used, ene reductases are available for this reaction. They require reduction equivalents in the form of reduced nicotinamide cofactors for their reaction. As ene reductases exhibit a broad substrate specificity, they have become popular among bioorganic chemists.

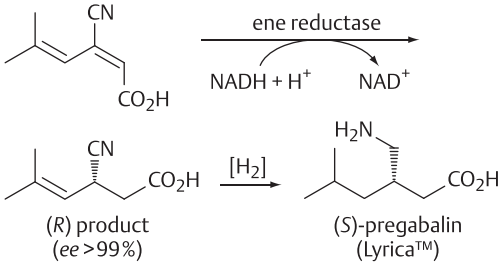
Chiral secondary amines by lipase-catalyzed amidation



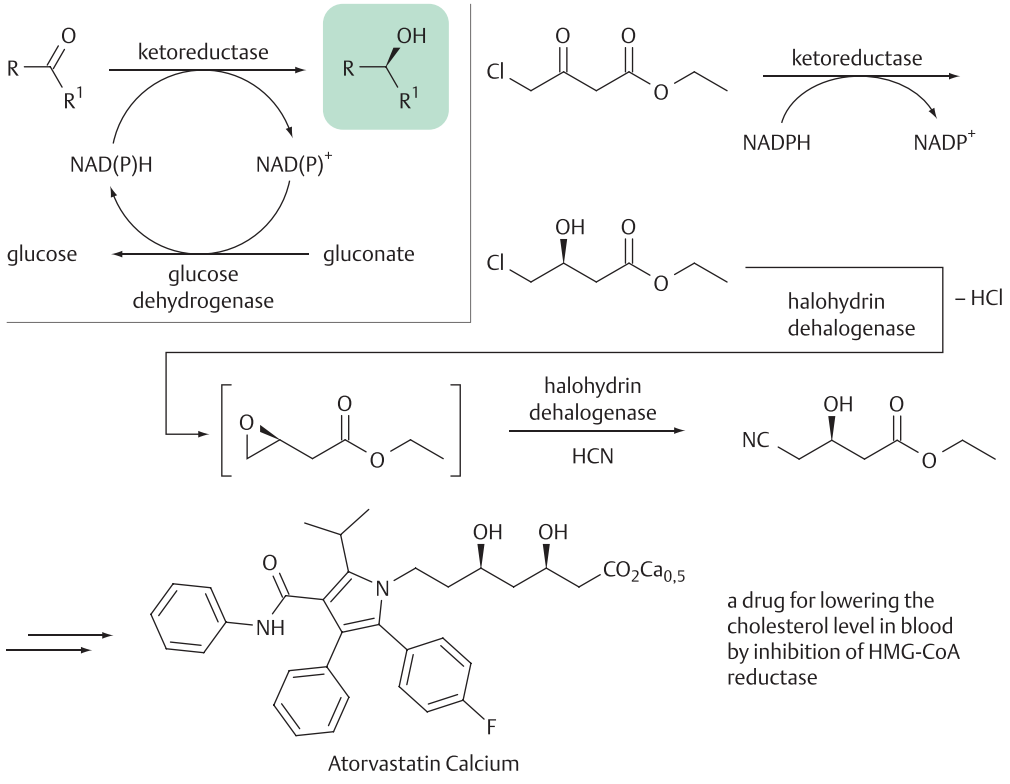
Selective hydratisation by nitrile hydratase



Asymmetric reduction of alkenes by ene reductases



Chiral secondary alcohols by reduction with ketoreductase



Enzymes as processing aids

General. Today, enzymes are used as processing aids in many areas of technology, e. g., in detergents (→174), in food technology, and in paper (→184), textile, and leather (→194) processing. Enzymes are also increasingly used for the industrial synthesis of fine chemicals, since they are often superior to chemical catalysts in regio- and stereoselectivity. For the same reason, enzymes are widely used in the analytical determination of medical samples and food products (→256). Many enzymes are available today in a recombinant form. This implies a lower price, higher purity, and the potential to optimize their properties for the desired use by application of protein engineering techniques (→198). These modern technologies are limited, however, to non-food applications, since food products containing components produced by genetic engineering techniques must follow declaration (→334) rules that might reduce consumers' acceptance (→336). For enzymes to be used as processing aids, hydrolases and isomerases are preferred since they do not require the addition of a cofactor.

Purposes. In industry, the use of any enzyme as an additive should eventually pay off. Improvements in product quality, savings in process costs, or environmental benefits are typical examples where this has proved true. For example, proteases in detergents contribute to dissolving protein stains from within fibers – an effect that cannot be had with chemical additives. The enzymatic hydrolysis of starch is superior to acid hydrolysis in terms of the formation of undesired byproducts. By using pectinases in fruit processing, the yield of juice can be significantly enhanced, and process costs during filtration are reduced. The use of glutaminases allows improved texture of meat, sausage, fish and soy products by creating new isopeptide bonds within the proteins contained in these food products. Proteases and collagenases permit selective removal of hairs and other skin components during leather treatment; their first application, about a century ago, dramatically improved the working conditions of tanners, who profession used to be outcast due to its distasteful working conditions. The coagulation of milk upon addition of microbial or recombinant rennet is significantly less expensive and more hygienic than the classical process, which was based on extracts from calf stomachs. In most of these cases, however, enzyme mixtures are used, partially for cost reasons and partially because side activities are a desirable fringe benefit (e. g., due to the presence of starch-degrading α-amylase in detergent proteases). As a result, the preparation of such mixtures is simpler and less demanding than that of analytical enzymes. On the other hand, characterization of the resulting enzyme mixtures is often difficult and is subject to a manufacturer's standards. Similarly, the technical or economic benefit of an industrial enzyme becomes apparent only if very specific determination methods are used, pertinent to a special craft or even a single manufacturer. Such methods may have developed over long periods of time and cannot be easily standardized, since internal quality control may rest on decade- or even century-long standards based on these individual methods. As a result, biochemical tests that are performed in an enzyme manufacturer's laboratory for screening improved enzymes may not match the application testing used in the client industry, often rendering the improvement of enzymes as additives very tedious.

Registration. (→334) The manufacture of enzymes as additives is regulated by the rules of GMP (good manufacturing practice). Enzymes produced by genetic or protein engineering techniques play a key role in leading to the higher ecoefficiency of detergents production. In smaller market segments, such as food enzymes, recombinant enzymes are the exception, not the rule. Though there may be a clear advantage for using a recombinant enzyme, the costly registration and real or expected consumers' concerns may form an insurmountable barrier to their use, except in some special cases such as recombinant chymosin (→188).

Enzyme additive costs and markets. The benefit of adding an enzyme preparation obviously relates to its price. Usually, the technical benefit translates into an economic advantage on the order of cents per kg product. As a result, the price for technical enzymes is mostly in the range of 3–10 € per kg. In spite of these limitations, the world market for technical enzymes has grown in 2013 to ca. 4–5 billion US$, and hydrolases used in detergents, followed by food and feed enzymes, have the largest share.

Enzymes as additives in industry

application	enzyme type	organisms (examples)	market size (% of total)	economic advantage
detergents	proteases, amylases, mannanases, cellulases, pectinases, lipases	*Bacillus licheniformis* *Aspergillus nidulans* *Trichoderma reesei*	26	1, 7, 8
starch hydrolysis	α-amylases, glucoamylases	*Bacillus amyloliquefaciens*	16	3, 4
glucose isomerization	glucose isomerase	*Streptomyces olivochromogenes*	5	1, 3
beer brewing	amylase	*Bacillus subtilis*	3	3, 4
fruit juice, wine	pectinases, xylanases, cellulases	*Aspergillus niger*	3	3, 4, 5, 6
baking	α-amylase, maltogenic amylase, xylanases, protease	*Bacillus amyloliquefaciens*, *Aspergillus oryzae*	8	1, 3
dairy (cheese)	proteases (renning), lipases	animal rennin, *Rhizomucor miehei*	8	2
animal feed	phytases, xylanases, proteases	*Aspergillus niger*, *Trichoderm reesei*	22	3
paper and textiles	cellulases, α-amylase	*Trichoderma reesei*, *Bacillus amyloliquefaciens*	5	4
leather treatment	proteases	*Aspergillus oryzae*	2	1, 7

process/application	enzyme cost per unit quantity (US$) (estimate)
starch liquefaction	ca. 10 per t starch
glucose from starch	3.5 per t starch
isomerization of glucose	6 per t starch
HFS in USA	6 – 7 per t starch
ethanol	1 per t starch
beer	0.1per 100 L
bakery goods USA	0.1per 100 kg flour
bakery goods EU	0.1 – 0.5 per 100 kg flour
fruit juice	0.1 – 0.5 per 100 L juice
wine	0.1 – 0.5 per 100 L wine
stabilization of fruit lemonade by glucose oxidase	0.3 – 0.8 per 1 000 L
cheese manufacture	0.05 per 100 L milk
detergents	0.05–0.25 per kg detergent
leather tanning	1.2 – 3 per t skin

important goals in application technology

1 higher product quality

2 improved taste

3 better yields

4 reduced process costs

5 better filtration

6 better conservation

7 improved working conditions, reduced environmental load

Detergent enzymes

General. About a century ago, Otto Roehm in Germany was the first to introduce detergents to which pancreatic enzymes had been added, to enhance the removal of proteinaceous stains such as blood, eggs, cocoa, grass stains, etc. When microbial proteases from Bacillus strains became available around 1960, the use of enzymes began to increase rapidly, and at present there is rarely a detergent without enzyme additives. Cloning and protein engineering (→198) have resulted in improved proteases that are perfectly adapted to laundry conditions. Recombinant Bacillus proteases (→20) are produced on a scale >10,000 t per year (calculated as poor enzyme protein). Other enzymes such as cellulases, lipases, amylases, hemicellulases, mannanases and pectate lyases are also used in laundry detergents.

Detergents and the laundering process. Laundry detergents usually contain anionic and nonionic surfactants, and sometimes cationic ones. Anionic surfactants are inactivated by the Ca^{2+} and Mg^{2+} ions of hard water; as a consequence, detergents contain complexing agents to sequester these ions. Instead of sodium pentasodium triphosphate, which used to be the most important complexing agent, sodium aluminum silicates in the presence of small amounts of "carriers" such as citric acid or phosphonates are preferred today. The traditional bleaching chemicals were sodium hypochlorite or sodium metaborate perhydrate; they have also been replaced by sodium percarbonate or medium-chain organic peroxi-acids. The pH of a detergent solution is ~10; the temperature during laundry washing is 30–90 °C and the time required is usually ~30 min. As a result, detergent enzymes must be adequately stable against alkali, surfactants, complexing agents, and bleaches. Further, their substrate specificity should be low.

Proteases. At present, only serine proteases obtained from Bacillus strains (subtilisins) (→20) are used. All strains are recombinant, with several copies of the protease gene behind a strong promoter (→62). Using protein design methods, detergent proteases have become more stable to complexing agents and oxidants. For example, the replacement of methionine[222] by alanine led to an enzyme with enhanced properties in dishwashers. In the fermentation process, production of the protease is usually initiated after completion of growth by addition of the promoter-specific inducer and is complete after 72 h or less. After complete removal of the cell mass by separators or membrane technology (→104), the extracellular enzyme is concentrated by precipitation or ultrafiltration, followed by partial purification. Since repeated inhalation of enzyme dust or aerosols, e. g., in working areas, may lead to allergic reactions, the enzyme concentrate is further processed into granules that are encapsulated and dust-free. For this purpose, the concentrate is sprayed on a core particle and then granulated in a rapid mixer in the presence of additives such as salts, waxes, and stabilizers, or it is directly sprayed on the core particle in the presence of the additives in a spray dryer. As a final step, each variety is coated with wax and pigment. In several large-scale studies, no allergies from inhalation or skin exposure of such products were found.

Cellulases. During the washing of cotton, endocellulases (→182) having a suitable pH optimum hydrolyze those cellulose microfibers that protrude from the fabric ("depilling"). As a result, the washed textile feels softer and its color is brighter. In addition, it resists soil pigments better than normal fibers. In detergents, cellulases from *Humicola insolens, Melanocorpus* sp. and *Thielavia* sp. that have been cloned into *Aspergillus oryzae* or *Bacillus subtilis.*are mostly used.**Lipases.** Lipases having an alkaline pH optimum are intended for removing long-chain wax esters (e. g., lipstick stains) that are difficult to emulsify with chemical surfactants. The most important detergent lipase presently used is from *Humicola insolens*, prepared in recombinant strains of *Aspergillus oryzae*.

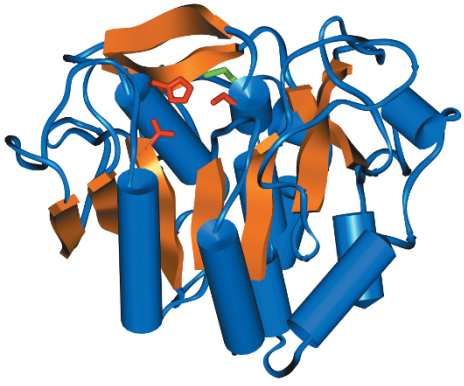
Amylases. Amylases (→176) loosen starch-containing stains by hydrolytic attack. They are used in concentrations up to 0.1 % amylase protein in automatic dishwashers. In essentially all washing and dishwashing powders, the amylases are alkaline- and thermostable and originate from strains such as *Bacillus stearothermophilus*, Thermoactinomyces or Pseudomonas. They may be modified by protein design (→198) for higher stability, and are expressed and industrially produced in recombinant *Bacillus amyloliquefaciens* or *Bacillus subtilis* strains.

Enzymes for detergents and automatic dishwashers

Enzyme	acts on	produced by	applied in*
protease	egg, milk, cacao, blood, gras	Bacillus (subtilisins)	W, D
amylase	starch, gravy	Bacillus (α-amylase)	W, D
lipase	skin fats, waxes (lipstick)	Humicola lipase (expressed in *Aspergillus oryzae*)	W
cellulase	dirt particles, pigments, cotton fiber (color refreshment, antipilling)	Aspergillus, Humicola and others (e.g., expressed in *Aspergillus oryzae*)	W
mannanase	guar gum (dressings, icecream)	Bacillus or Trichoderma	W, D
pectinase	marmelades, fruits	Aspergillus	W, D

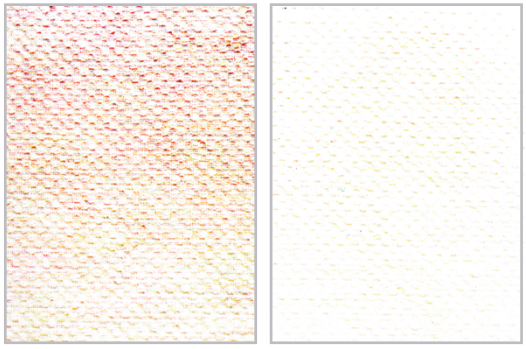
W = washing machine, D = automatic dishwasher

Structure



subtilisin Carlsberg (1sbc) at 0.23 nm. Red: catalytic triad (ser, asp, his); green: met^{222}

Detergency of protease



heavy duty detergent without protease
60 °C washing cycle

heavy duty detergent with protease
60 °C washing cycle

Fermentation and recovery

preculture

genetically modified high-performance strain of *Bacillus lentus;* 10 m^3, 24 h at 35 °C

bioreactor

120 m^3, 48 h at 35 °C C-source: dextrin; N-source: soy meal, induction by casein

>25 g L^{-1} subtilisin in 60 h

cell separation and concentration

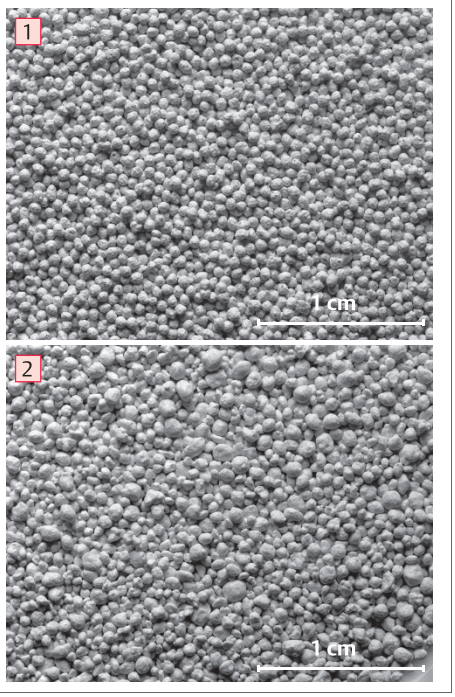
microfiltration and ultrafiltration or cross-flow filtration, precipitations

1 layered granulate

core particle (coated salt/sugar granulate), radical scavengers, PEG, TiO_2, layered coating in spray dryer

2 mixer granulate

core particle (salt/PEG), formed to granulate in mixer, dryed, coated by PEG/TiO_2 in spray dryer



Enzymes for starch hydrolysis

General. Starch is, after cellulose, the second most important polysaccharide produced on earth. Next to glucose, it is the most important carbon source for fermentation (→328). In 2012, ca. 75 million t of starch were produced; 70 % came from corn, followed by potatoes and cassava. Only ca. 20 % of isolated starch is used directly; 30 % is chemically modified, and 50 % is saccharified to yield glucose or its oligomers, the dextrins and maltodextrins. The preferred method of hydrolyzing starch is with enzymes, which leads to fewer side-reactions than acid hydrolysis.

Starch is a polymer (dp 200–5000) that is composed of linear amylose (poly-α-1,4-D-glucose) and branched amylopectin. Amylose is pseudocrystalline. Its amount in starch can be determined by the iodine-starch reaction. In the amylopectin component, the linear amylose chain is branched at about every 20 glucose residues, in a D-1,6-glycosidic manner. Depending on the origin of the starch, the ratio of amylose to amylopectin may vary; this ratio determines the physical and chemical properties of the starch. Starch is insoluble in cold water. When heated, it dissolves to the extent to which intramolecular hydrogen bridges are destroyed. In the gelatinization range, starch forms a gel by absorption of water, resulting in a strong increase in viscosity. In this state, it can be chemically modified or enzymatically degraded. If gelled starch is cooled, amylose recrystallizes quickly with the formation of intermolecular hydrogen bonds (retrogra-dation). Starch is an important component or raw material for many basic food materials such as bread or tapioca. Traditional procedures for their manufacture are being gradually replaced by protocols in which enzymatic processing of starch plays an important role.

Starch-degrading enzymes. Several enzymes are available for degrading starch: 1) α-amylases (synonym: exo-amylase) hydrolyzes starch at α-1,4 positions within the polymer chain; 2) β-amylase (synonym: endo-amylase) hydrolyzes maltose or maltotriose from the non-reducing end; 3) glucoamylase (synonym: γ-amylase, amyloglucosidase) splits maltose to two moles of glucose, but also, if with reduced velocity, the α-1,6-bonds of amylopectin; 4) pullulanases split α-1,6-bonds, preferentially of pullulan, but also of amylopectin; and 5) isoamylases also split α-1,6-bonds, but their velocity is higher with amylopectin than with pullulan.

α-Amylases. This aspartyl enzyme occurs in many organisms. It has been crystallized from malt, pancreas, and *Aspergillus oryzae* (→16). Crystal structures have been obtained for the α-amy-lases of *Bacillus amyloliquefaciens* (→20) and other organisms. α-Amylases have been cloned and overexpressed in various host systems. Bacterial α-amylases exhibit higher temperature optima (*Bacillus licheniformis*: 78 °C) and are also more alkali-stable than fungal amylases. As a consequence, a wide range of α-amylases of different pH- and temperature stabilities are available for any desired application.

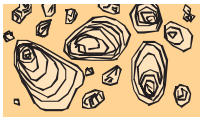
β-Amylase, a sulfhydryl enzyme, is available from wheat malt and also from *Bacillus stearothermophilus*. It is an important enzyme in the preparation of maltose syrups.

Amylases splitting α-1,6-bonds. The most important enzyme of this group is glucoamylase from *Aspergillus niger*. A similar enzyme from *Rhizopus* sp. is also used. Pullulanase is industrially produced by strains of *Klebsiella pneumoniae* or *Bacillus cereus*.

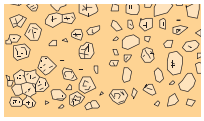
Manufacturing procedures. For applications in food products, amylases are produced using GRAS strains and sometimes still in solid state fermentation (SST) (→86). Amylases for technical applications (bioethanol, syrups, detergents, textiles), which are often temperature-stable enzymes, are mostly cloned into engineered *Bacillus amyloliquefaciens* or *Bacillus licheniformis* production strains that are optimized for performance and stability. For production, aerated tank fermenters from 30–200 m^3 are used. Harvest is after 36–48 hrs, and amylase concentrations are in the order of 15–30 g/L. The enzymes are secreted into the culture medium, resulting in a dilute enzyme solution. Since the enzyme product is of relatively low value, the isolation procedures include only a few simple process steps starting with removal of the cell mass by separators or rotatory drum filters (→104), followed by concentration by, e. g., ultrafiltration, precipitation, and finishing, in the presence of stabilizing additives. The market value of temperature-stable α-amylases and glucoamylases for bioethanol production is estimated at 400 million US$ (2012).

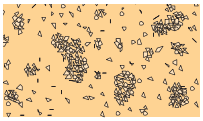
Composition and properties of various starches

starch from	amylose (%)	amylopectin (%)	gelatinizing range (°C)	swelling capacity (fold)
potato	18 – 23	77 – 82	56 – 66	> 1000
wheat	19 – 25	75 – 81	52 – 63	21
corn	21 – 30	70 – 79	62 – 72	24
rice	17 – 19	81 – 83	61 – 78	19

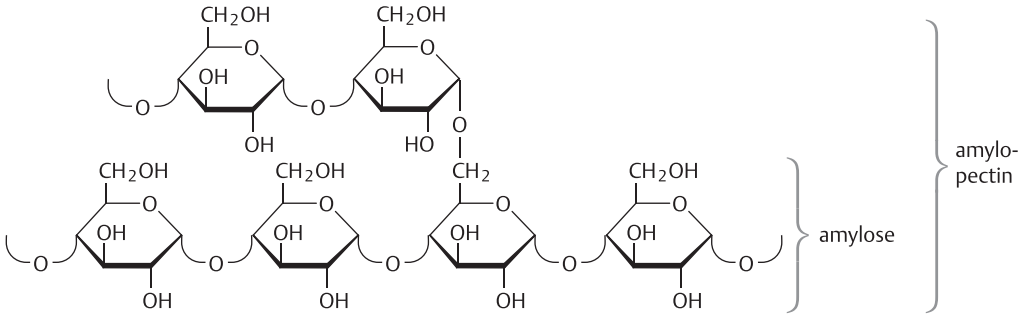

potato


wheat

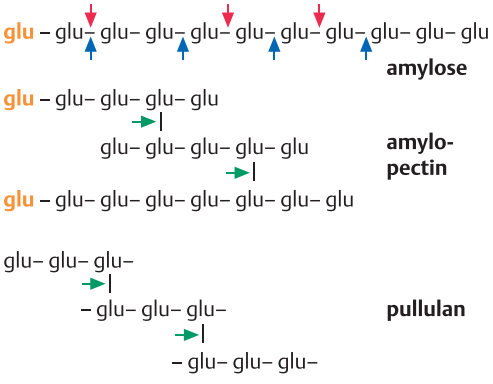

corn


rice

Structure of amylose and amylopectin



Enzymatic degradation



- α-amylase (1,4-α-D-glucan-4-glucanohydrolase, EC 3.2.1.1): splits α-1,4-bonds to maltose and maltotriose
- β-amylase (1,4-α-D-glucan-galtohydrolase, EC 3.2.1.2): splits maltose units from the non-reducing ends
- glucoamylase (γ-amylase, (1,4-α-D-glucan-glucohydrolase, EC 3.2.1.3) and pullulanase (amylopectin-6-glucanohydrolase, EC 3.2.1.41): splits α-1,6-bonds

mixture of α-amylase and glucoamylase/pullulanase: splits amylose and amylopectin to D-glucose

glu = nonreducing end

Enzymes

enzyme	origin	properties	conditions of application
α-amylase	*Bacillus licheniformis*	best effect at 60 °C and pH 5.7, requires Ca^{2+}	95 – 105 °C , pH 6 – 7, Ca^{2+}
	Aspergillus oryzae	best effect at 50 – 60 °C and pH 5.0, requires Ca^{2+}	< 50 °C , pH > 3.5, Ca^{2+}
	barley malt	best effect at 70 °C and pH 5.5	< 70 °C , pH > 4.5
β-amylase	*Bacillus stearothermophillus*	best stability at 75 °C, pH 5.0	for maltose syrups
pullulanase	*Klebsiella pneumoniae*	best effect at 50 – 55 °C and pH 5.0	55 – 65 °C , pH 3.5 – 5
gluco-amylase	*Aspergillus niger, Trichoderma* reesei	best stability at 55 – 60 °C and pH 4.0	55 – 65 °C , pH 3.5 – 5

Enzymatic starch hydrolysis

General. About half of the starch that is isolated annually (ca. 75 million t y^{-1}) is enzymatically hydrolyzed. In the USA, most hydrolyzed starch is used in the manufacture of bioethanol or isoglucose (also termed HFS: high fructose syrup) (→168, 180). The remainder is partially hydrolyzed to dextrins and maltose syrups, which are used in a wide range of applications, e. g., as a carbon source in fermentations (→328). In the USA, corn and wheat are mostly used as raw materials for starch. Potato and rice starch have less practical importance as raw materials for glucose, since they are produced by intensive agriculture in Europe or Asia and thus cannot compete in price. When Napoleon blocked the access of British merchandise to Europe (the Continental Blockade), Kirchhoff developed a method in St. Petersburg to prepare sugar from potato starch with dilute sulfuric acid as a replacement of imported cane sugar. This technology, however, leads to colored by-products and today is unattractive.

Enzymatic starch hydrolysis. Starch from corn or wheat is obtained by wet milling. Valuable by-products are corn germ oil, corn and wheat gluten, and feed additives of various compositions. The milled starch is heated for only a few minutes, in the presence of thermostable bacterial α-amylase, to temperatures between 105–140 °C ("starch cooking"), to swell and gelatinize the starch. After 2–3 h in the presence of Bacillus α-amylase, biotransformation to maltodextrin (dextrose equivalent DE 15–20) is complete. Maltodextrin is a mixture of oligosaccharides with minor quantities of mono-, di-, and trisaccharides. It is an excellent starting material for the ensuing complete saccharification. Maltodextrins are also used as food components of low sweetness, e. g., in baby food, hospital diets, and instant soups.

Enzymatic saccharification can be directed to yield dextrose, glucose, high-maltose, or high-conversion syrups. In this process, glucoamylases (→176) from *Aspergillus niger*, *Aspergillus oryzae* or *Trichoderma reesei* (→16) are used. They are either used as native enzymes or they are produced from the native strains after self-cloning and protein engineering (→198), resulting in improved performance and higher yields. In the process, maltodextrin obtained from the starch cooker is cooled to ca. 60 °C and the pH is adjusted to ~4. Usually, the process is carried out continuously, requiring several tanks in sequence to prevent mixing syrups of different degrees of hydrolysis. This procedure results in glucose syrups with a very high DE value (97–98) after 48–72 h. Addition of pullulanase or isoamylase (→176) leads to even higher DE values and reduces the glucoamylase levels. The use of immobilized enzymes has proved of little value, due to diffusion limitation in the highly viscous substrate solution and the formation of reversion products. From syrups with very high DE, pure D-glucose monohydrate (dextrose hydrate) can be isolated by crystallization. Using an appropriate choice of enzymes and saccharification conditions, a wide choice of glucose syrups of different compositions can be tailor-made. Glucose syrups with relatively low and medium DE are used in the manufacture of sweets. High-maltose and high-conversion syrups are often produced from maltodextrin using α-amylase from *Aspergillus niger*; they have high viscosity, but a reduced tendency to form brown colors or to crystallize.

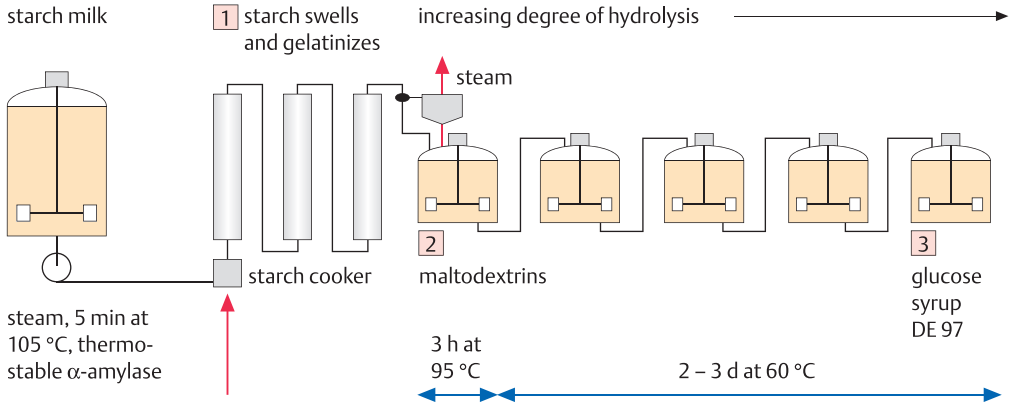
Cyclodextrins. If maltodextrins are processed with the enzyme cyclodextrin transferase, 5-, 6-, or 7-membered cyclodextrin rings are formed. They exhibit good solubility in water, but also contain hydrophobic cavities 0.5–0.75 nm in diameter, which can accommodate hydrophobic guest molecules such as vitamins, fragrances, malodors or drugs. In consequence, cyclodextrins are used to enhance the solubility of such compounds in formulations, increase their bioavailability, improve their stability, or add or remove flavor compounds. Due to their surface-active properties, cyclodextrin can also be used as an remulsifier, for example in mayonnaise. All cyclodextrins have obtained GRAS status from the Food and Drug Administration, and α-cyclodextrin has a verified health claim in the EU for reducing blood sugar peaks. Chiral cyclodextrin derivatives are also used as stationary phases for the separation of mixtures of chiral compounds. The industrial manufacture of cyclodextrins is mostly limited to β-cyclodextrin; suitable cyclodextrin transferases for this process have been isolated from mesophilic and alkalophilic Bacilli. Starch or dextrins are used as a starting material in this process.

Enzymatic degradation of starch

1 corn starch hydrolysis
40 % w/v, deionized water, 20 ppm Ca^{2+}, pH 6 – 6.5, 77 – 105 °C, 0,15 % bacterial α-amylase

2 maltodextrins DE 15 – 20
2 – 3 h at 95 °C , bacterial α-amylase

3 glucose syrup DE 97
pH 4.2 , 60 °C, 2 – 3 d, mold glucoamylase

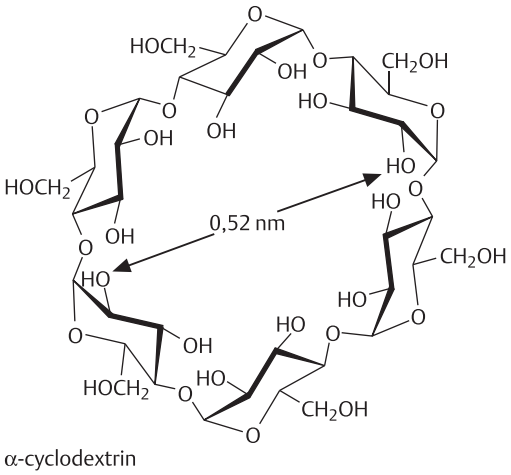


Products of enzymatic starch hydrolysis

product	characteristics	enzyme used	applications
maltodextrins	DE 15 – 25	α-amylase	food additive with good rheological properties; raw material for sweeteners
maltose syrups	DE 40 – 45	α-amylase, β-amylase	sweetener
high maltose	DE 50 – 55	α-amylase, glucoamylase	sweetener
maltose syrups with high or extra high degrees of saccharification	DE 60 – 70 DE < 80	α-amylase, glucoamylase, pullulanase	sweetener raw material for fermentation
dextrose	DE 97	prolonged hydrolysis	raw material for isoglucose
isoglucose	DE 97	glucose isomerase	sweetener

DE = dextrose equivalent, a measure for the degree of starch hydrolysis

Cyclodextrins



α-cyclodextrin

Properties

cyclodextrins	α	β	γ
number of glucose residues	6	7	8
M_R	972	1135	1297
water solubility (in 100 g/mL)	14.5	1.85	23.2
diameter (nm)	≥ 0.47	≥ 0.6	≥ 0.75
CAS	10016-20-3	7585-39-9	17465-86-0

Enzymes and sweeteners

General. Sugar (D-saccharose, sucrose) became an important food additive in Western culture only in the 18th century. It was originally isolated from sugarcane grown in tropical or subtropical countries. In response to Napoleon's continental embargo of Europe, the production of sugar from sugar beets developed as a new technology. Today, corn and wheat starch have become very important raw materials for sugar production: in two enzymatic processes, starch is first hydrolyzed to glucose, followed by isomerization to glucose-fructose syrups (isoglucose). Due to dietetic considerations (reduced intake of calories, cavity prevention), several sugar substitutes have been developed. The sweetness of sugars or sugar substitutes is evaluated, using 10 % solutions, by a sensory panel in comparison to the sweetness of saccharose solution in water. High intensity sweeteners such as Thaumatin or Advantam may be over 1000-fold sweeter than saccharose. The food and beverage industry is increasingly replacing sugar or corn syrup with artificial sweeteners in many products that traditionally contain sugar.

Invert sugar. Saccharose is hydrolyzed by acid or by invertase to invert sugar syrup, an equimolar mixture of D-glucose and D-fructose. The sweetness of invert sugar is similar to that of saccharose. Since it does not crystallize, it is often used in the manufacture of candy and soft chocolate fillings. Invertase is obtained from the cell mass of bakers' yeast, *Saccharomyces cerevisiae* (→120). It is located in the cytoplasm and can be obtained in pure form, after breakage of the cell wall in a ball mill, by a small number of downstream processing steps. Invert sugar is produced in an enzyme reactor (→102) with immobilized invertase, using 70 % saccharose syrup as the starting material. The stability of the enzyme permits the reactor to be used continuously for several days at 55 °C.

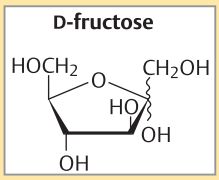
Isoglucose. The enzyme xylose isomerase (systematic name: D-xylose-aldose-ketose isomerase, EC 5.3.1.5), which can be isolated from Streptomyces and other strains, isomerizes in an equilibrium reaction not only D-xylose to D-xylulose, but also D-glucose to D-fructose. At the preferred reaction temperature of 60 °C, a mixture of 42 % D-fructose and 55 % D-glucose is formed (isoglucose, high-fructose syrup HFS) which is slightly less sweet than saccharose. If glucose is separated from this mixture by chromatography and re-isomerized, isoglucose 55 can be obtained. It contains 55 % D-fructose. Its higher sweetness has proven advantageous in the manufacture of soft drinks. By far the most important producer of isoglucose 42 and 55 is the USA. In 2012, US isoglucose production was 7.5 million t, 37 % of the caloric sugar market. Production is slightly decreasing, however, due to the higher use of sugar substitutes. Due to EU regulations protecting sugar beet farmers, only 690,000 t of isoglucose may be produced in the EU (ca. 5 % of the beet sugar production (2013)).

Manufacture. The intracellular enzyme glucose isomerase occurs in several microorganisms. This protein forms a homodimer, and bivalent metals, e. g., Mn^{2+} or Co^{2+}, participate in catalysis. About 1500 t of this enzyme is produced annually, mostly from Streptomyces strains or *Bacillus coagulans*, usually in a fed-batch bioreactor process (→192) taking 72 h. To avoid the cost of isolating the intracellular enzyme, isoglucose is often manufactured directly with the immobilized microbial strain, which is inactivated, crosslinked by glutaraldehyde and bound to a suitable carrier material (→102). Within 1 h of contact time between the immobilized microorganism and the glucose feed, the concentrations of substrate and product reach equilibrium. At a process temperature of 60 °C, this is at ~42 % D-fructose, yielding isoglucose 42. The half life time of the catalyst under these conditions is around 50 d at 60 °C. Thus, continuous operation of the process is possible if a modular process line is used. Due to strain optimization, modern processes no longer depend on the addition of Mn^{2+} or Co^{2+} as a cofactor. The reaction product contains traces of caramelized sugars and is purified by passing it over charcoal columns. Isoglucose is marketed as a syrup.

D-Fructose is a sugar of high sweetness. It can be prepared from invert sugar or from isoglucose by chromatographic procedures, usually by simulated moving-bed chromatography. Topinambur (Jerusalem artichoke) tubers contain up to 75 % inulin, a fructose polymer (→118). Fructose can be prepared from topinambur in an enzymatic process using inulinase. An enzyme from *Aspergillus niger* (→16) is mostly used for this process.

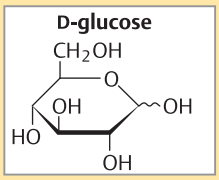
Sugars used as sweeteners

D-fructose



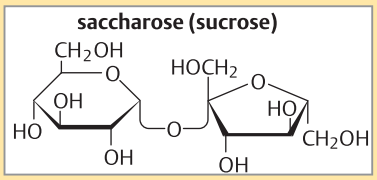
M_R 180.16
m. p. 106 °C, decomp.

D-glucose



M_R 180.16
m. p. 146 °C, decomp.

saccharose (sucrose)



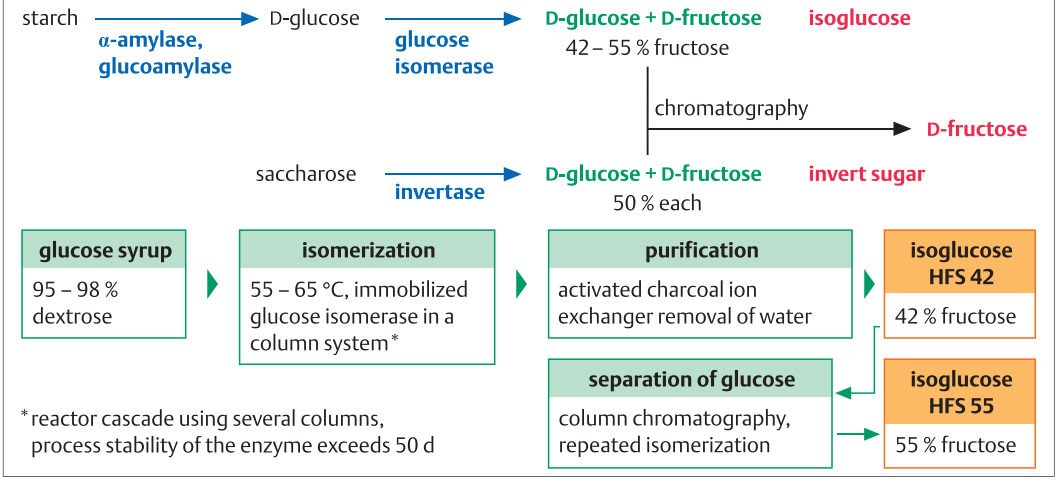
M_R 342.30
m. p. 185 –186 °C, decomp. above 165 °C

Sugars and sugar substitutes

name	relative sweetness (related to saccharose)	starting material, production process
saccharose	1.00	isolated from sugar cane or sugar beets
glucose	**0.5 – 0.8**	**hydrolysis of starch with α-amylase, glucoamylase**
glucose syrups	**0.3 – 0.5**	**hydrolysis of starch with α-amylase, glucoamylase**
hydrogenated glucose syrups	0.3 – 0.8	hydrogenation of starch hydrolysates or glucose syrups
isoglucose 42	**0.8 – 0.9**	**enzymatic isomerization of glucose with glucose isomerase**
fructose	**1.1 – 1.7**	• **enzymatic hydrolysis of saccharose, or** • **enzymatic isomerization of glucose, or** • **enzymatic hydrolysis of inulins** followed by chromatographic separation
invert sugar	1.00	**hydrolysis of saccharose with invertase**
mannitol	0.4 – 0.5	hydrogenation of fructose
sorbitol	0.4 – 0.5	hydrogenation of glucose
xylitol	1.0	hydrogenation of xylose
lactitol	0.3	hydrogenation of lactose
maltitol	ca. 0.9	hydrogenation of maltose
palatinitol, isomaltol	**0.45**	**enzymatic isomerization of saccharose to isomaltulose (palatinose), follwed by hydrogenation yielding a mixture of glucopyranoside sorbitol and glucopyranoside mannitol**

the relative sweetness of a sweetener is determined organoleptically by a panel and compared to a 10 % aqueous solution of saccharose. As a result, values can fluctuate around a median

Manufacture



*reactor cascade using several columns, process stability of the enzyme exceeds 50 d

Enzymes for the hydrolysis of cellulose and polyoses

General. Cellulases and hemicellulases are used in many areas of biotechnology. Their use for the enzymatic hydrolysis of biomass is currently being studied intensely, in view of the generation of glucose as a fermentation raw material and of energy (usually bio-ethanol) (→138, 328). For such processes, biomass must be rendered by physical or chemical processes into a soluble form, to allow for the access of the enzyme catalyst. Such processes are considered essentially viable but have not yet become economically feasible. Cellulases and xylanases were expressed in or on the surface of organisms such as *Corynebacterium glutamicum* (→20) or *S. cerevisiae* (→14), to allow for the production of products such as amino acids or bioethanol from dissolved biomass. Cellulases are also used in the food industry for the gentle manufacturing of vegetable and fruit purees (→186). Other applications include pulp- and paper-manufacturing processes (→184). In some detergents, alkaline cellulases are used as softening agents for cotton fibers and textiles (depilling) (→174).

Cellulose is the shape-determining material of the plant cell wall. It is also the most abundant renewable resource: the annual production of cellulose in the biosphere is estimated to be ca. 20 billion t. Cellulose is built of β-1,4-linked units of D-glucose (average degree of polymerization ca. 10,000), which are ordered via hydrogen bonds into parallel bundles (microfibers).

Polyoses (hemicelluloses) are a heterogeneous group of heteropolymers built from pentoses, hexoses, deoxyhexoses, and hexuronic acids. They constitute ca. 20 % of the cell wall. Xyloglucans (xylans) are directly linked to the cellulose microfibers via hydrogen bonds. They consist of a backbone of β-1,4-linked glucose residues that carry numerous xylose residues as β-1,6-branches, which can reach a considerable length. Arabinogalactans bind to the glycoproteins of the cell wall; their backbone is made up of alternating galactose and arabinose residues linked by β-1,4-bonds. Pentosans consist of alternating residues of xylose and arabinose.

Biodegradation. In nature, celluloses and hemicelluloses are broken down by bacteria and especially by white-rot fungi (→16). These fungi hydrolyze cellulases and hemicellulases in parallel with lignin oxidation, using a number of unusual enzymes. The process starts with mechanical softening of the cell wall via growth of the fungal mycelium. In some organisms, e. g., the Clostridia, a number of cellulose-degrading enzymes exhibiting various activities occur as an ensemble combined on a platform, the cellulosome.

Cellulases can be isolated from many microorganisms. Among the better known bacteria are Cellulomonas and Clostridium strains. Fungal cellulases can be isolated from *Trichoderma reesei, Aspergillus niger, Humicola insolens*, and others. The manufacture of these enzymes in a bioreactor follows the usual pattern for extracellular enzymes. The enzyme preparations obtained after cell separation and fractionated precipitation may still contain side activities from hemicellulases, which may be desirable for the intended application. Cellulases for the saccharification of wood are mainly prepared from the white-rot fungus *Trichoderma reesei*. A mixture of cellulolytic enzymes may be manufactured by surface or submersed fermentation in a bioreactor. If surface cultures are used, the enzyme mixture is obtained by extracting the koji with water. In submersed fermentations, the cells are removed and the broth is fractionated by the addition of ethanol.

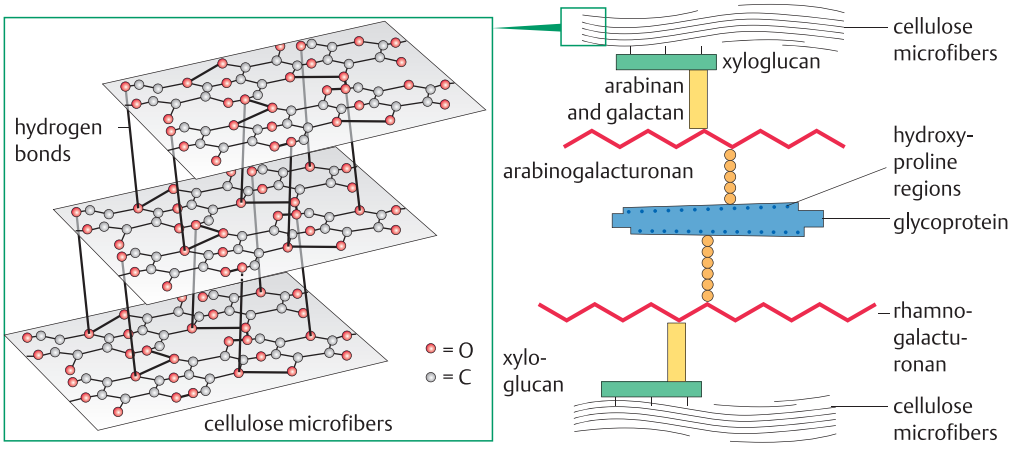
Hemicellulases. β-Glucanases are prepared by isolation from *Bacillus subtilis, Penicillium emesonii, Aspergillus niger*, and other strains. Mannanases and galactomannanases are isolated from *Aspergillus niger* or *Trichoderma reesei*. Manufacture follows the standard rules for obtaining extracellular microbial enzymes.

Glucose and xylose. For any production methods that start with cellulose or biomass (wood chips, bagasse, and others), the costs of transporting the substrate to the bioreactor and preparing it for fermentation, as well as the cost of the enzyme, are the key parameters. In substrate preparation, lignin is mostly removed by heat and/or aggressive chemical treatments, which are similar to the those used in preparing pulp. In this process, xyloses are partially degraded. Enzyme costs have become greatly reduced through the development of high-performance recombinant production strains that secrete cellulases or hemicellulases.

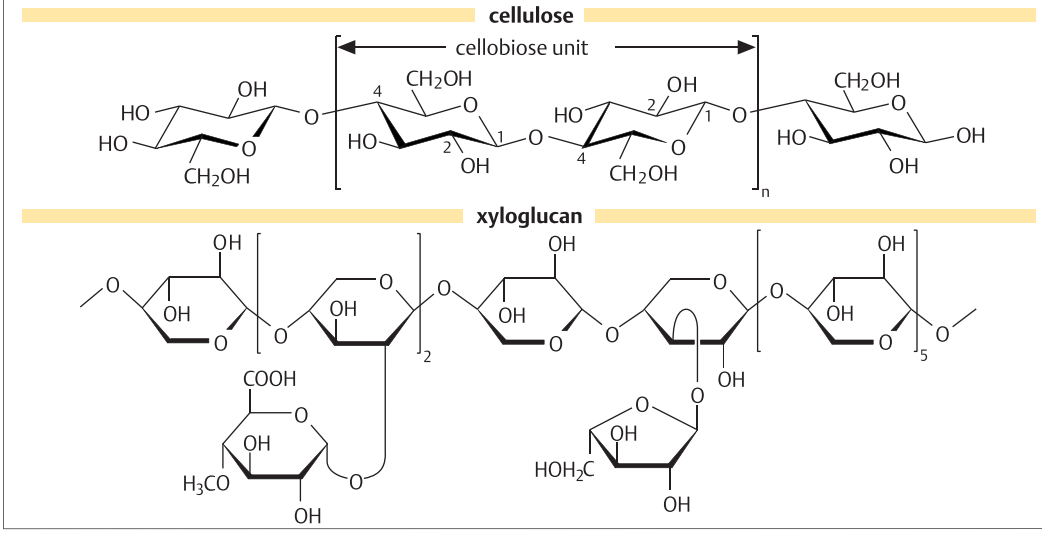
Composition of polysaccharides in plant cell walls

component	structure	degree of polymerization	building block	percentage
cellulose	β-1,4-D-glucose	1,000 – 10,000 microfibers	D-glucose	wood*: 60 % cotton: 90 %
pectins	polygalacturonic acids, rhamnogalacturonic acids, galactans, arabinogalactans	100 – 2 000	D-galacturonic acid, D-galacturonic acid methyl esters, L-rhamnose, D-arabinose	10 – 40 %
polyoses (hemicellulose)	xylans, xyloglucans, β-1,3- and -1,4-D-glucans, galactomannans, arabinogalactans, glucuronomannans		xylose, glucose, galactose, mannose, arabinose, glucuronic acid	20 %

*wood contains ~ 40 % lignin as second major component



Cellulose and xylan



Manufacture of cellulase (example)

seed culture	bioreactor	recovery	
recombinant strains of *Trichoderma reesei*	150 m^3, 7 – 14 days, <100 g/L active enzyme protein	filtration or separator, spray drying or ultrafiltration	46% of dry mass is active enzyme protein (AEP)

Enzymes in pulp and paper processing

General. Pulp is a fibrous material that consists mainly of cellulose. It is produced by removing most of the lignin and hemicelluloses from wood. Pulp is manufactured into paper, cardboard, and chemical products. In 2011, ca. 180 million t of pulp and ca. 400 million t of paper and cardboard were produced (25 % of the total amount in China, 22 % in North America). For producing 1 t of paper, 3.3 t of wood are required; in addition, 0.4 t of petroleum is needed for energy. The wastewater from pulp and paper plants is highly polluted; it contains 1 kg AOX and 55 kg COD per t of pulp. As a consequence, improved manufacturing processes are being studied; they should require less material and energy and be less harmful to the environment. The properties of the end products pulp and paper depend both on the type of tree and on the manufacturing process. Besides some hardwoods that grow exceptionally fast (eucalypti, poplars), softwoods are preferred (birches, pines, firs), since their wood consists of longer fibers, which are better to process. Relative to hardwoods, softwoods contain less polyoses (14–17 %) but more lignin 26–32 %). Research based on plant tissue culture and genetic engineering is ongoing in an attempt to reduce the lignin content of softwoods (→284).

Pulp manufacturing. After the trees are cut and their bark peeled, the wood is mechanically cut into chips, which are further transformed into pulp by mechanical, thermomechanical, or chemical means. The dominant chemical process today is the Kraft process (80 % of world production), which is based on the alkaline depolymerization of lignin using Na_2S/NaOH under high pressure at 170 °C. Alternatively, lignin can be removed by sulfonation, treating wood with an excess of sulfite (sulfite pulp). In both processes, a bleaching step using chlorine or chlorine derivatives follows. Enzyme-based process steps are being studied in the context of biopulping and for a milder bleaching of the raw pulp.

Biopulping. Prior to their mechanical degradation, pretreatment of the wood chips with lignin-degrading microorganisms is being studied (biopulping), mostly using white-rot fungi (→16). For example, in the Cartapip™ process of Clariant, wood chips are sprayed with nutrients and spores of the white-rot fungus *Ophiostoma piliferum* and are left for some weeks. This process was investigated to a scale of 100 t; when followed by a Kraft process, yields were higher and the quality of the products was improved.

Enzymatic bleach (biobleaching). The pulp produced both by Kraft and sulfite treatment is of a loose structure and easily accessible to an enzyme. The subsequent bleaching with ClO_2 is necessary to remove the colored by-products of the pulping that mostly originates from lignin. The pulping process can be improved and less ClO_2 is required if the pulp is pretreated with xylanases and/or laccases, especially if it comes from hardwoods. Several paper mills have explored this enzymatic procedure up to a production scale. The amount of ClO_2 could be reduced by 5 kg t^{-1} pulp, compensating for the price of the enzyme; and the wastewater load was reduced by a third. Xylanase (→182) complexes from *Clostridium thermocellum* and *Streptomyces roseiscleroticus* were studied in this process, as well as a xylanase from the deep-sea thermophile *Thermotoga maritima* (temperature optimum 96 °C) expressed in *E. coli* as a host. Examples of commercially available enzymes are the xylanases from *Trichoderma longibrachiatum* and *T. reesei*. The application of these technologies is, however, still under development.

Pitch control. Some woods, e. g., pine, contain a rather high amount of triglycerides which, under the drastic conditions of pulp manufacture, tend to form pitches that later deposit on the paper machines and impair the whiteness of paper. Microbial lipases have proved valuable for pretreating pine woodchips to prevent pitch formation. Another benefit of this method is that no chlorinated fatty acid material is formed during chlorine bleaching. At present, this may be the most important application of any enzyme in pulp and paper processing. Laccases, e. g., from *Trametes versicolor*, are also studied in combination with a mediator to remove pitches originating from sterols and lignin phenols.

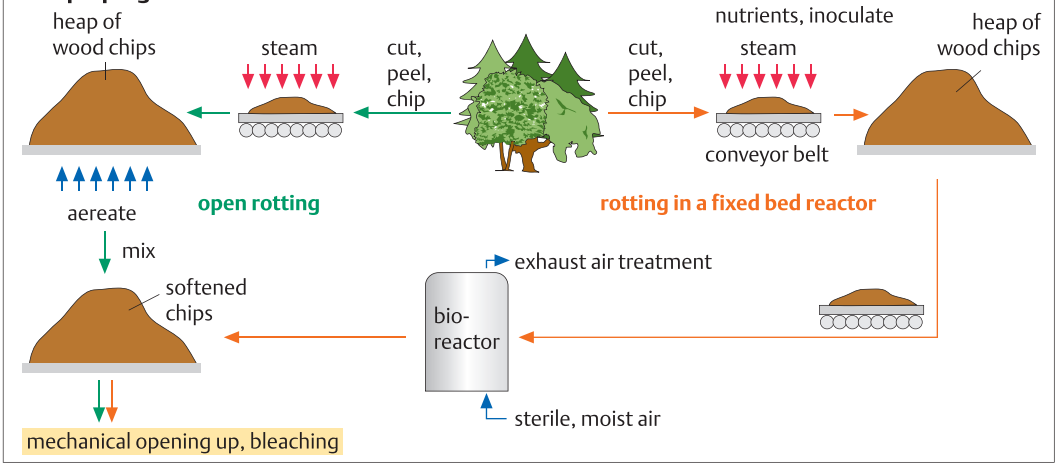
Removal of printers' ink. In recycling printed paper, printers' ink is removed more thoroughly if the paper is pretreated with a mixture of cellulase, xylanase, and lipase.

Processes for pulp and paper manufacture

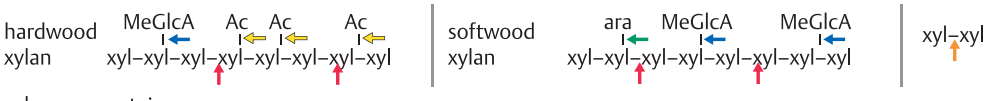
wood chips → mechanical pulping → raw pulp → bleach → pulp → additives → paper

method	conditions	chemical processes	remarks
sulfate method (Kraft process)	Na_2S, NaOH 2 h, ~170 °C pH 13 – 14	lignin degradation by HS^-	~ 80 % of all papermills; cellulose, polyoses are attacked as well, resulting in only partial opening up
sulfite method	SO_2 , HSO_3^-, SO_3^{2-} depending on the (Ca, Mg, Na, NH_4^+) pH 1 – 13 1 – 4 h, 125 – 180 °C	lignin degradation by sulfonation	compared to Kraft process, lignin is less well degraded
chlorine bleach	Cl_2, OCl^-, ClO_2	lignin degradation and bleaching of colored byproducts by chlorination	formation of ecologically doubtful compounds
O_2^-, H_2O_2-bleach		lignin degradation and bleaching of colored byproducts by peroxidation	

Biopulping



Enzymatic degradation of xylan and pulp bleaching

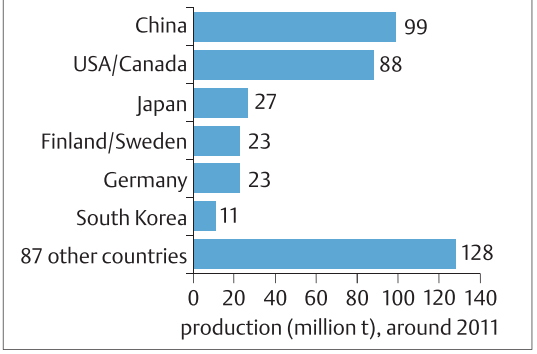


xylanases contain:

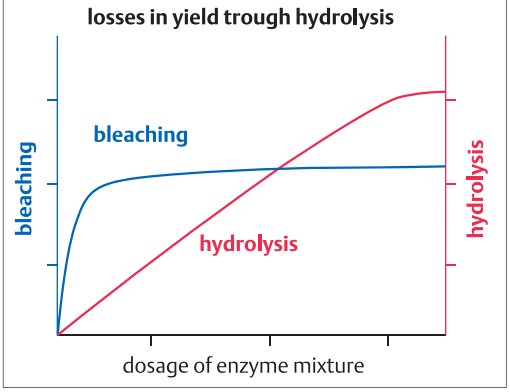
→ β-xylanase → α-glucuronidase → α-arabinosidase → esterase → β-xylosidase

xyl = xylose, ara = arabinose, MeGlcA = 4-*O*-methyl-D-glucuronic acid, Ac = acetyl

Paper and cardboard manufacture



losses in yield trough hydrolysis



Pectinases

General. Technical pectinases are usually composed of a variety of different hydrolases and lyases. They degrade pectin, thus modulating the texture and viscosity of mashed fruit and vegetables while preserving natural colors. As a result, they are important enzymes in fruit and vegetable processing and are manufactured on a scale of ca. 1000 t y^{-1} (world).

Pectins are acidic polysaccharides with a M_W 30,000–300,000 Da. Their composition varies widely, but their main building block is δ-1,4-linked D-galacturonic acid, which may be esterified with methanol. D-Rhamnose-rich sequences may also occur, carrying sidechains of D-arabinose, D-xylose, and D-galactose. Pectins of high molecular weight that are highly esterified form the water-insoluble middle lamella of the primary cell wall of plants (protopectin, "cell concrete"). They may constitute up to 2 % in fruit juices and more than 20 % in fruit pulps (e. g., 25 % in lemon pulp). Through depolymerization and hydrolysis of the ester bond, protopectin is transformed into shorter-chain pectins that carry an anionic charge and thus can form cross links and gels in the presence of Ca^{2+} ions. This process, which occurs naturally during the ripening of fruits and vegetables, leads to major changes in structural, physical, and chemical properties. Thus, the mild degradation of pectins by pectinases is in demand during the industrial manufacture of purees and fruit juices.

Pectinases include the following enzymes:

1. endopolygalacturonases [EC 3.2. 1. 15]
2. exopolygalacturonases [EC 3.2. 1. 67]
3. pectate lyase [EC 4.2.2.2]
4. exopectate lyase [EC 4.2.2.9]
5. pectin lyase [EC 4.2. 2. 10], and
6. pectin esterase [EC 3.1. 1. 11].

Commercially available pectinases contain different ratios of the above enzymes, leading to the desired property during application.

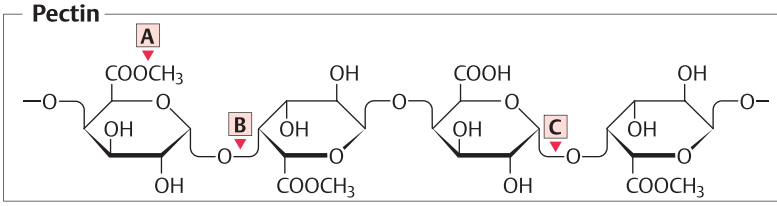
Technical preparation. Presently, all commercial pectinases are prepared from molds such as Aspergillus or Rhizopus whose enzymes have been cleared for food applications (GRAS status) (→166). Many preparations are still prepared by surface fermentation (→86) of these molds. For example, the medium may contain wheat bran held in high humidity. After ca. 100 h cultivation, the culture is extracted with buffer solution. Ultrafiltration of the extract leads to an enzyme concentrate that is microfiltered to remove spores. After the addition of stabilizers such as glycerol or sorbitol, the preparation is standardized as to its enzyme activity (which may be quite difficult, because of the presence of a wide range of enzyme activities, e. g., cellulases, hemicellulases, and glycosidases) and sold on the market. In addition to concentrates, spray-dried enzyme preparations are also available. Preparations with well-defined activities could be prepared by genetic engineering techniques, but their acceptance is still doubtful, due to consumers' concerns (→336).

Applications. Pectinase preparations are used 1) for maceration of vegetables and fruits ("macerases"), 2) for treating mashes during preparation of fruit juices or grape juice, to enhance the filterability and yield, 3) in the treatment of grape must, for removing suspended pectin material, and 4) in processing citrus fruits, mango or pineapple to prevent the formation of gels. The macerating action of pectinases is often used in the manufacture of fruit and vegetable purees for baby foods, turbid fruit juices, and fruit yogurts. If pectinase is added during juice manufacture, yields can be improved by 5–10 %, and filtration is enhanced by a factor of 1.5–5, depending on the type of fruit. During wine production, pectinases are used, for example, for maceration, improved extraction and juice clarification, and for faster filtration.

Cellulases and hemicellulases (→182) are often used in food processing. Examples include the preparation of starch from corn and the extraction of coffee beans and tea leaves. If added to pectinases, they can further improve vegetable and food processing. Bacterial or fungal glucanases help to open up the malt used in breweries, thus rendering faster mash formation during the beer production process (→112). In combination with food-grade proteases, cellulases are used to keep ultrafiltration membranes clean and permeable. Arabinases from *Bacillus subtilis* and xylanases from Aspergillus, Trichoderma or Bacillus strains have found applications in the clarification of juices such as pear and apple juice.

Phospholipases are used for the "degumming" of plant oils by removing phospholipids that act as emulsifiers and reduce processing yields.

Pectin



A pectin esterase
B pectin lyase*
C pectate lyase*/poly-galacturonase

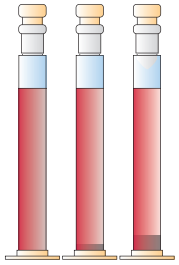
*form anhydro sugars

Juice production (e.g., grape, red currant, apple juice)

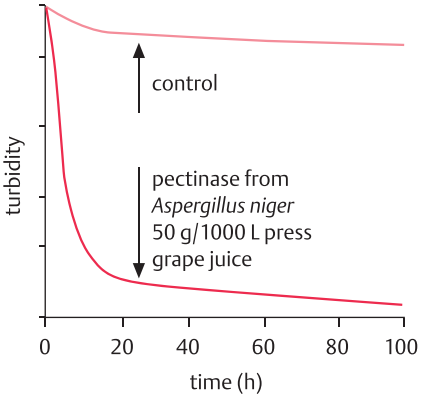
berries
wash, stem, pit, peel, mill, crush, chop

mash
pectinase preparations, possibly cellulases/hemicellulases, e.g., 30 min at 30 °C, press, decant, strain

turbid juice
pasteurize or treat with pectinase, filter through membrane

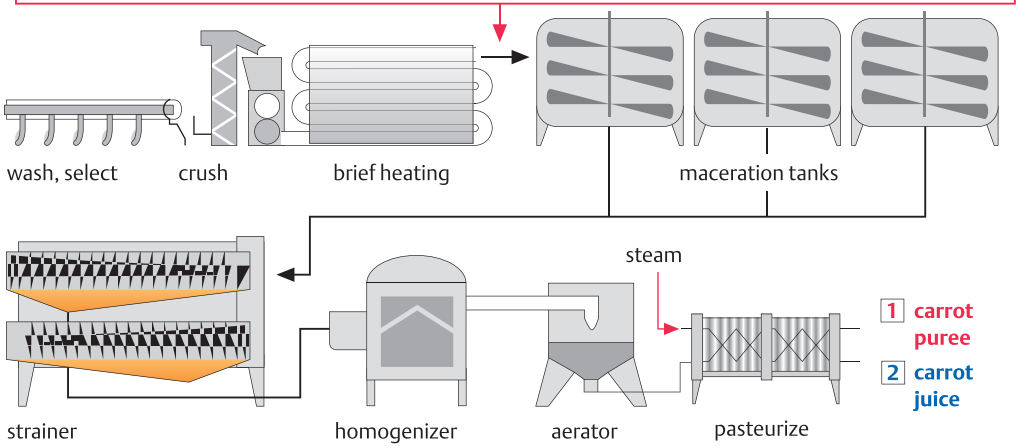


pectinase from *Aspergillus niger*

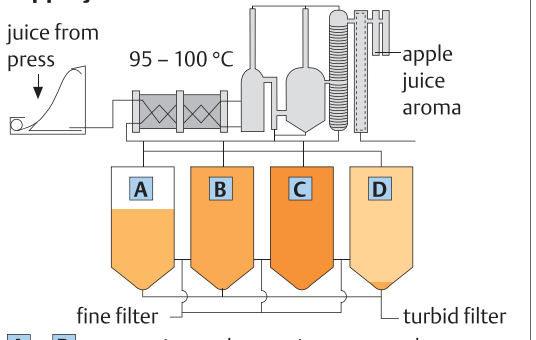


Processing of carrots

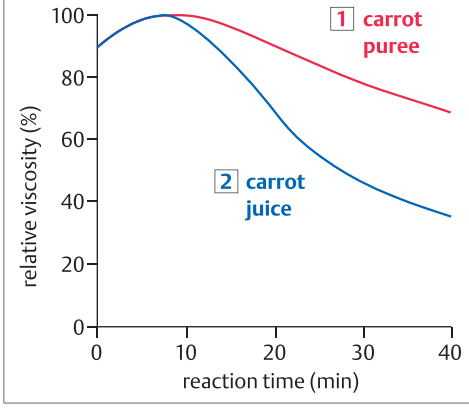
1 pectinase from *Aspergillus niger* 2 pectinase and C1-cellulase from *Aspergillus niger*



Apple juice concentrate



A + B enzymation tanks: pectinase, α-amylase
C + D clearing tanks: gelatin, bentonite



Enzymes and milk products

General. During the manufacture of milk products, proteases, lactases, and lipases are mainly used. A key enzymatic process is the use of highly specific proteases (rennins) for 18 million t of cheese production (world); some 1000 t of enzyme are used for this purpose. Secondly, the aroma of cheese can be influenced by the action of lipases and proteases. Since lactose (milk sugar) is not tolerated well by many adults, lactase (β-galactosidase)-treated milk products have been developed. In some developing countries, milk products are stabilized against microbial attack by adding lysozyme or H_2O_2/catalase. Finally, whey has been investigated as a fermentation raw material; it is produced in quantities of 2 million t (2011, world) as dry powder during the preparation of cottage cheeses and cheese.

Milk is an oil-in-water emulsion with a water content of ca. 90 %. The triglycerides (→34) of milk (butterfat) are unusual in that they contain 2–4 % butyric acid. The percentages of lactose and protein are in the range of 3 % each. Casein, the main component of the protein fraction, is a mixture of phosphoproteins with M_R 20–30 kDa. It forms aggregates of ca. 1 million Da, the κ-casein fraction serving as a protective colloid. To become digestible, this colloid must first be degraded. This can occur in various ways: by the presence of Ca^{2+} ions in concentrations >6 mM, by a decrease of the pH to values <4.6, and through hydrolysis of a single peptide bond of κ-casein (^{105}phe–^{106}met) with chymosin (rennin), a protease of the upper digestive tract of mammals.

Casein-hydrolyzing proteases. In the stomach of mammals, casein is first hydrolyzed by the combined action of chymosin and acid, becoming susceptible to further proteolytic attack. The related technical procedure for transforming milk into sour milk or cheese by the addition of cow, goat, or sheep rennet is one of the oldest inventions for preserving this highly perishable food. In traditional cheese making, preparations of animal stomachs were used for this purpose. They were, however, often in short supply, rather impure and difficult to standardize. As a consequence, today microbial rennets or, more recently, recombinant chymosin is preferred. Microbial rennet, a protease from *Mucor miehei*, exhibits the same specificity (^{105}phe–^{106}met) as animal chymosin and is widely used. In 1987, calf chymosin was cloned and expressed in *E. coli*. Presently, *Saccharomyces cerevisiae*, *Aspergillus niger* (→16) or *Kluyveromyces lactis* are the preferred host organisms. Since 1992, the use of recombinant chymosin in cheese making has been allowed by the FDA in the USA, and in the EU and Japan since 1997. Whereas in Central Europe there is still skepticism from consumers (→336), in the USA, the UK and other countries most cheese is produced using recombinant chymosin.

Lactose hydrolysis. Lactose (milk sugar) is tolerated well by most mammals only until maturity. Adult humans, with the notable exception of most Northern Caucasians, exhibit a similar intolerance towards lactose. This occurs because the formation of lactase becomes largely reduced by adulthood, resulting in the passage of undigested lactose into the large intestine; there it is fermented by the intestinal flora, leading to excessive gas formation. In primary lactose intolerance, this is caused by a single locus (T/T at position 13910: no intolerance, C/C: intolerance). Lactose intolerance must be clearly distinguished from the rare genetic disease galactosemia. The latter is due to an autosomal recessive gene defect on chromosome 9, leading to a phenotype in which UDP-galactose is not synthesized. This results in an overproduction of galactose metabolites, in particular of toxic galactol. Whereas a galactose-free diet is indispensable for people who suffer from galactosemia, lactose intolerance can be avoided by not ingesting lactose or by enzymatic hydrolysis of lactose in milk products. Milk products pretreated with lactase are available for the manufacture of fermented milk products, e. g., yogurts (→116), for hydrolyzed lactose syrups used in bakeries, and for animal feeds based on whey. The sweetness of hydrolyzed lactose exceeds that of the native disaccharide, adding sweetness to lactase-treated products.

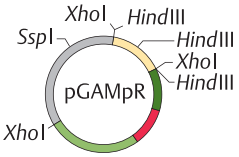
Cheese aroma. The short- and medium-chain fatty acids occurring in butterfat can be partially hydrolyzed by lipases to a product mixture useful for their aroma in cheese production (enzyme-modified cheeses, EMC). Depending on the chain-length specificity of the lipase used, various aroma notes can be produced.

Composition of milk

	milk (%)	whey (%)
water	~ 88	~ 94
fat	~ 3 – 4	~ 0.5
protein	~ 3.3	~ 1
casein	~ 2.6	–
lactose	–	~ 4.8

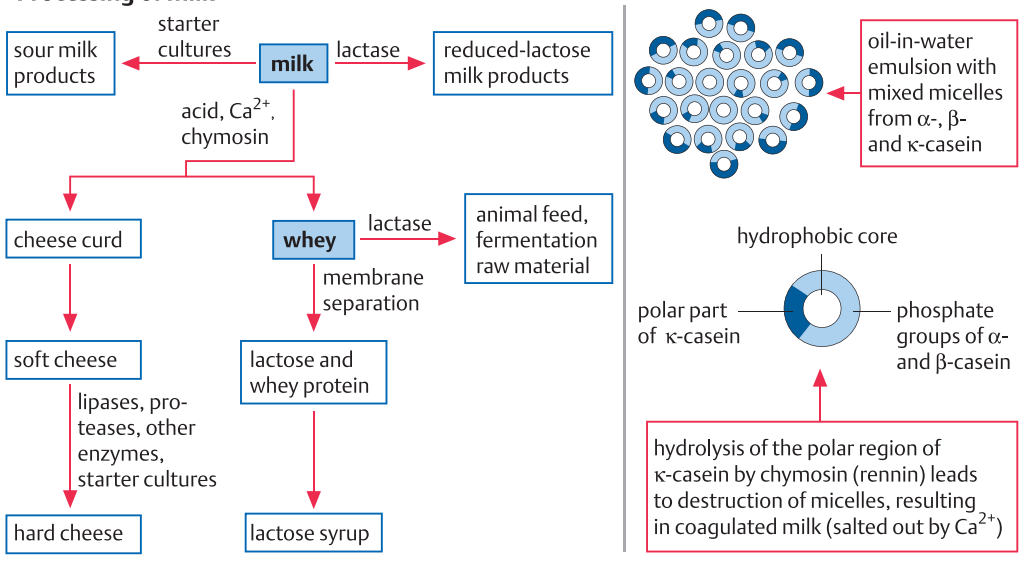
Plasmid for the expression of prochymosin

in *Aspergillus niger* var. *awamori*



- *Neurospora crassa* pyr4 gene
- glucoamylase A promoter- and coding sequence
- preprochymosin DNA
- glucoamylase A terminator sequence

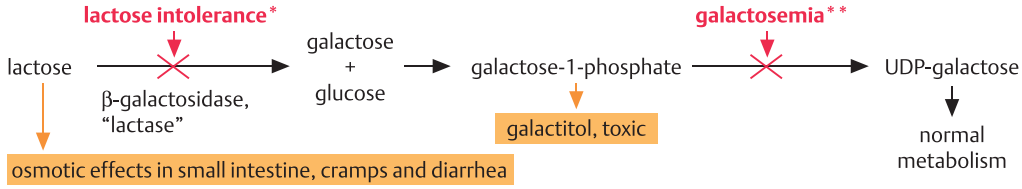
Processing of milk



Manufacture of chymosin

native	microbial	recombinant
stomachs of young animals	**preculture**	**recombinant microorganism**
cutting, activation at pH < 5	high-yield mutants of *Mucor miehei* or *M. pusillus*	*Escherichia coli*
extraction	**bioreactor**	**bioreactor**
salt water, 14 d	dextrose syrup, soy meal, 30 °C, 72 h	maltodextrins, 37 °C, 36 h
purification	**purification**	**purification**
ultrafiltration standardization	separation of mycelium, reverse osmosis, precipitation	isolation of inclusion bodies, Tritron-X100/EDTA, urea-/alkali-extract, ion-exchange chromatography, acid treatment
200 U/kg stomach	5 000 U/m^3 in 72 h	20 000 U/m^3 in 36 h

Lactose intolerance and galactosemia



* > 70 % of adult Bantus, American Blacks, Indians, Chinese, Aboriginees

** galactose-1-phosphate-uridyltransferase defect on chromosome 9, frequency 1:100 000

Enzymes in baking and meat processing

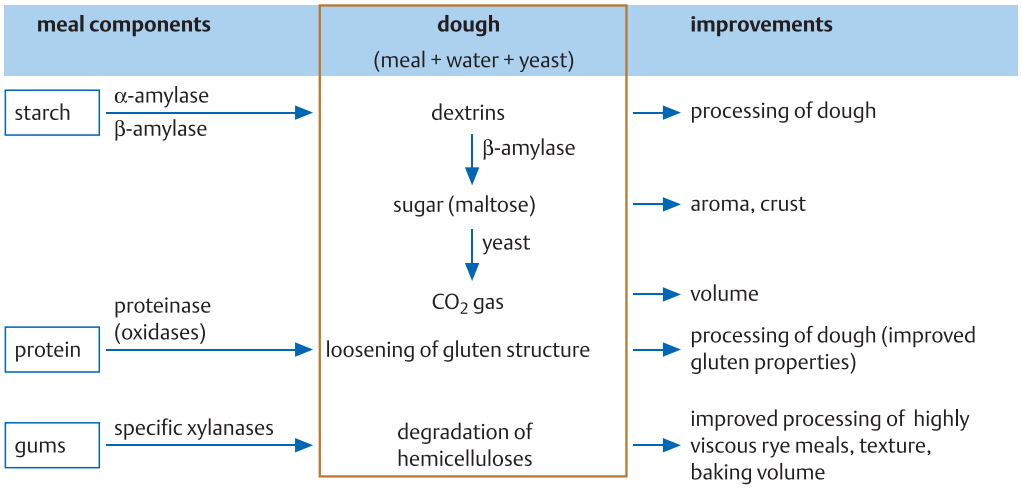
General. The most important applications of biotechnology for bakery products are the preparation of yeast dough and of sourdough. For sourdough, a combined fermentation of lactic acid bacteria and *Saccharomyces cerevisiae* increases the digestibility of rye meal. Enzymes such as amylases, xylanases, lipases, glucose oxidase and proteases are used as well in processing dough and in bakery goods. Specific xylanases are used to improve the formation of pentosane-based gels binding moisture and improving the dough processing of rye and wheat flours. Amylases, lipases, glucose oxidase and proteases can be advantageously used in many flour types since they are able to adjust the properties of starch and of gluten, leading to a desirable viscosity of the dough and an improvement in its fermentability. Maltogenic β-amylase (→176) is used as a processing aid to slow the staling rate of bread. The use of enzymes in the baking industry is estimated at around 1000 t (world). For the processing of sausages. starter cultures are often used (see Starter Cultures) (→114). They influence both aroma and preservation. Meat can be tenderized by the addition of proteases; papain and bromelein are major proteases used for this application.

Meal processing and enzymes. The milling of grain leads to a flour whose composition with respect to starch, pentosans and proteins depends strongly on the type of grain (wheat, rye, etc.), but also on the composition of the soil, the climate and the time of harvesting. The flour characteristics can be adjusted through the addition of appropriate enzymes. Amylases depolymerize the starch to dextrins (α-amylase), maltose (β-amylase) and finally to glucose (→176). As a consequence, their addition influences the processing, the aroma and the volume of dough, as baker's yeast can ferment only mono- and disaccharides. A maltogenic β-amylase from *Bacillus stearothermophilus* has proven an excellent processing aid to reduce the staling of bread. The gluten protein of the grain creates a protein network during the mixing of dough, forming a gel. If proteases are added. this protein matrix is selectively loosened and the extensibility of the dough is increased – a prerequisite for good retention of the CO_2 that is formed during fermentation. An excess addition of protease leads to a reduction of the viscoelastic propenies and thus to a destruction of the overall structure. In practice, amylase from malt, from Bacillus or fungal strains are used. Fungal glucoamylase is also applied (→176). Since the artisanal and industrial production of bread requires specific technological functions for mechanical dough processing, enzymes are widely used as processing aids. For proteases, the preferred enzyme has been isolated from *Aspergillus oryzae*. It loosens the gluten in a selective manner without leading to excessive degradation. As all enzymes remain in the products, recombinant enzymes are not used at present. Rather, enzyme mixtures produced by surface or tank fermentation of the producing microorganisms that contain side activities are isolated, characterized and sold in an appropriate matrix, containing stabilizers such as buffers.

Analytical methods. The field is dominated by traditional methods related to traditional craftsmanship. Thus, gas formation is monitored in rheofermentometers, and the decrease in viscosity is measured by a "falling number" method. For the evaluation of enzyme adjustments, farinographs, extensograms or alveograms are used, which all give some evidence about the viscoelastic properties of the dough.

Meat and enzymes. Meat is formed from muscle through complex biochemical transformations after the supply of oxygen has stopped. Proteases (cathepsins) play an important role in these transformations. From a consumer's point of view, a meat that is juicy, easy to chew, and of a pleasing texture, color, and taste are the major concerns. In Western cultures, this was traditionally achieved by storage and marinating. Other cultures wrap meat in papaya leaves or dip it into pineapple juice to tenderize it and enhance the aroma. Spraying meat with papain, a sulfhydryl protease isolated from papaya leaves, has the same effect. Alternatively, inactivated papain, in which the sulfhydryl group is oxidized with H_2O_2, can be injected into the bloodstream of animals shortly before they are killed. Once the animal has died, the sulfhydryl groups of the papain are reactivated after the oxygen in blood has been reduced, resulting in rapid partial hydrolysis of the meat.

Meal processing and enzymes

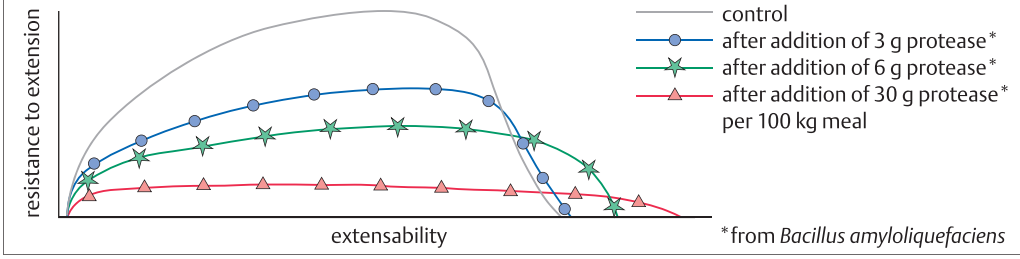


Enzymes for baking

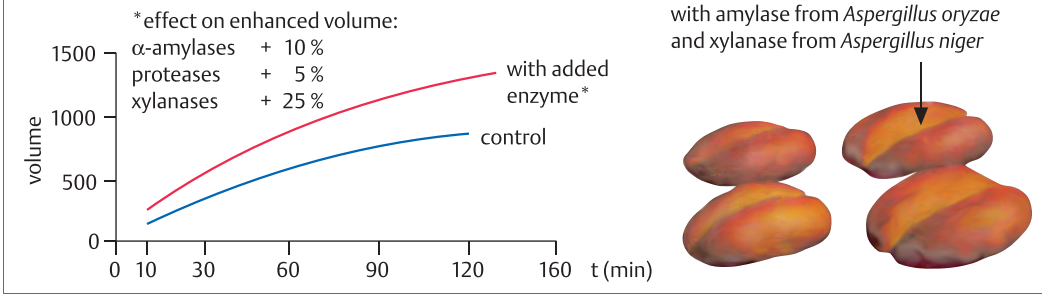
enzyme	effect on meal	method of determination	effect on baked product
α-amylase from malt, Bacillus, Aspergillus glucoamylase from *Aspergillus oryzae*	opening up of damaged starch grains and sticky starch to maltose and glucose	gas formation*, viscosimetry (amylograph or falling number), baking experiments	increased volume, better crust and taste
neutral proteases from molds	degradation of gluten	viscoelastic properties in farinograph, extensogram etc.	enhanced viscosity of dough, gas is better retained
xylanases from *Trichoderma viride*	hydrolysis of gums	viscosimetry in amylograph	improved fermenting tolerance
β-amylase from *B. stearothermophilus*, Aspergilllus	prevents interaction of starch and gluten	baking experiments, organoleptic test	prolonged freshness, better texture, antistaling

**Saccharomyces cerevisiae* cannot metabolize higher oligosaccharides, but only maltose and glucose

Extensability of dough



Baking experiment



Other enzymes for food products and animal feed

General. Several enzymes have recently found new practical applications. Transglutaminase is an enzyme that forms intra- or intermolecular bonds in proteins. It is used for the modification of protein-based foods such as meat or soy products. Asparaginase transforms protein-bound asparagine residues into aspartate, reducing the formation of carcinogenic acrylamide on heating. And feed enzymes contain phytase, cellulases and xylanases (→182). The latter enzymes improve digestibility of cellulosic materials in feed. Phytase is used for the processing of phytate, a major storage form of phosphorous (P) in seeds.

Transglutaminase [EC2.3. 2. 13] is an enzyme that modifies the structure of proteins by forming either intra- or intermolecular crosslinks through acyl group transfer from protein-bound glutamine residues to the ε-amino group of protein-bound lysine residues, forming an isopeptide-bond. Factor XIII of the blood coagulation system has transglutaminase activity: it forms isopeptide bonds during the solidification of a fibrin mesh ("retraction phase") (→228). These bonds confer high stability to proteins against proteolysis. A microbial enzyme from a GRAS organism *Streptomyces mobaraensis* was found to catalyze the same reaction. It is now widely used to modulate the texture and properties of protein-containing food. Thus, it is used for gelling concentrated solutions of proteins such as soybean proteins, milk proteins, beef, pork, chicken and fish gelatin and myosins. It is used, e. g., for the preparation of imitation crabmeat, for fishballs and restructured meat products. A commercial product is Activa TG™ from Ajinomoto. Transglutaminases are thiol enzymes, forming thioesters as intermediates that can be amidated by protein-bound lysine ε-amino groups.

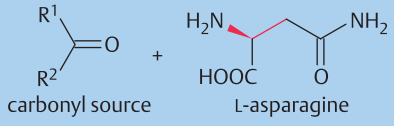
Asparaginase [EC3.5.1.1]. When starchy food is heated by baking, frying, toasting or roasting, some of the protein-bound L-asparagine residues are being transformed by a Maillard reaction into acrylamide, a potential carcinogen. The addition of L-asparaginase to processed food hydrolyzes some of the L-asparagine residues into L-aspartate, which does not form acrylamide upon heating. As food processing usually involves steps at elevated temperature, heat-stable L-asparaginases are required. Acrylaway HighT™ from Novozymes, one of the L-asparaginases marketed for this application, originates from *Pyrococcus furiosus*, a deep-sea anaerobic microbe.

Phytate and phytase. Phytates are salts of *myo*-inositol hexakisphosphate with potassium, magnesium and calcium. They form the major storage form of phosphorus in seeds and have functions beyond being a P-source for germination. Ruminants degrade phytate through phytases produced by their rumen microorganisms (microbiota), releasing phosphate. Monogastric animals such as swine and poultry do not produce phytase, and phytates are thus not available as a P-source. As a consequence, supplementation of phytase to grain-based feed – where up to 80 % of P is bound to phytate – enhances P-supply to monogastric animals, leading to faster growth. Their addition to feed also reduces phytate-bound phosphate in liquid manure that would lead to eutrophication of rivers and lakes. Phytases have been isolated from many plant and microbial sources, and x-ray structures of several phytases are available. They share a common sequence in their active sites (histidine acid phosphatases). Most phytases cleave 5 out of the 6 phosphate groups of phytic acid at different rates. According to EC nomenclature [EC3.1.3.n], they are grouped with reference to the position of the first hydrolyzed phosphate group; most phytases characterized so far are 3- or 4-phytases. As a feed additive, major phytases in commercial use are derived from thermophilic bacteria or fungal strains and expressed in, e. g., *Aspergillus oryzae* or *Pichia pastoris*. Thermostability is important as feedstuff is prepared by pelleting or other thermal processes. They are optimized by screening or protein engineering (→198) for fast release of phosphate and for transient temperature stabilities up to 95 °C which occur during feed processing. Phytases have also been expressed in transgenic plants and in transgenic pigs, in view of enhanced P supply.

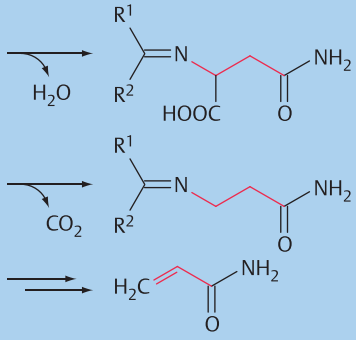
Cellulases and Xylanases. Cellulose and xylans (→182) are fiber components of cell walls in both viscous (wheat, rye, barley) and non-viscous cereals (corn and sorghum). The addition of cellulases and xylanases to cereal diets for monogastric animals like chicken or pigs leads to enhanced supply of energy for animal growth. Commercial xylanases (endo 1,4-ß-xylanase) are manufactured from strains such as *Trichoderma reesei* and are thermostable enough to withstand the high temperatures occurring during feed production.

Reduction of acrylamide: L-asparaginase

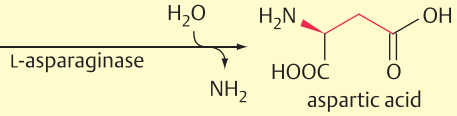
Formation of acrylamide



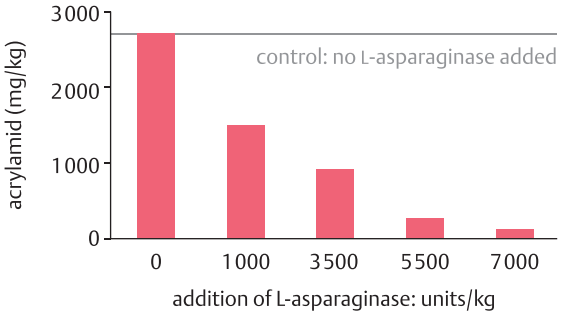
e.g., glucose



... and its prevention by a competitive reaction

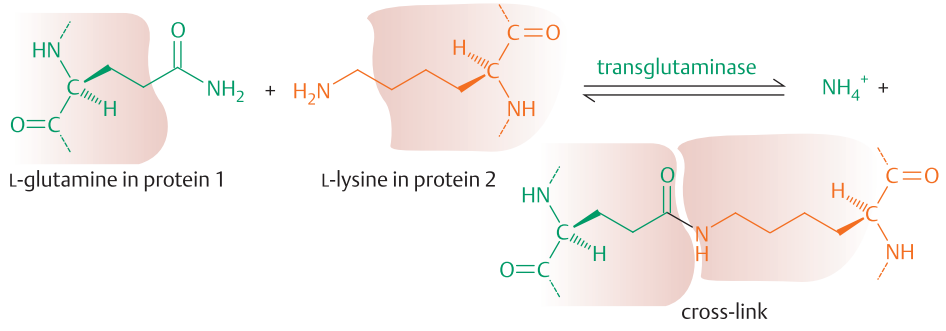


Potato tubes: frying time 15 sec

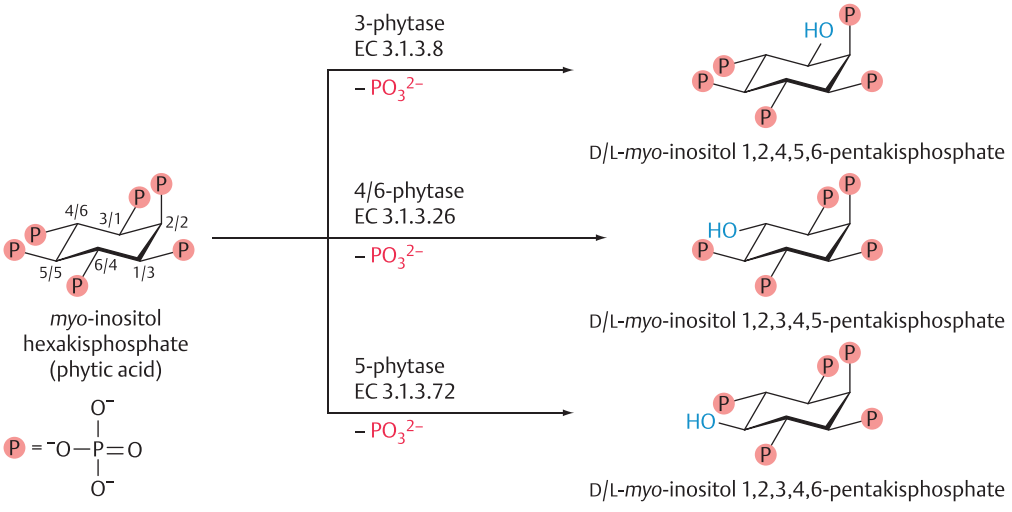


A trial demonstrating the amount of acrylamide remaining in potato snack pellet tubes after frying for 15 seconds. Samples were treated by different Novozymes Acrylaway® dosage.

Transglutaminase: a meat binder



Phytase enhances supply of phosphate



Enzymes in leather and textile treatment

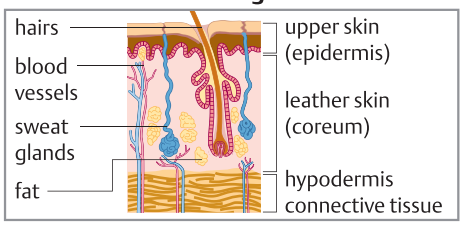
General. The preparation of leather from animal hides can be traced back to antiquity. Since rather harsh chemicals such as lime, alkali, and sulfur have been used throughout history, and feces and urine were unknowingly used as an enzyme source, tanning was considered an "unclean profession". In 2004, Otto Roehm in Germany was the first to lay the foundation for a science-based technical approach to leather treatment, resulting in a highly improved image of this craft. Proteases for leather treatment today are used at a level of several 100 t per year (world).

Leather. Animal skin consists of 60–65 % water, ca. 30 % proteins (>90 % collagen, with keratin, elastin, and others as minor components), and 2–10 % fat. The epidermis (upper skin) is just ca. 1 % of the skin; it is further divided into the outer *stratum corneum* (horny layer), a granular layer, and a mucosal layer. The dermis/*corneum* (leather skin) makes up 85 % of the skin's thickness; it is composed of the papillary layer, with collagen fibers, and the reticular layer, made up of connective tissue. The final layer is the hypodermis, which contributes about 15 % to the skin's thickness and contains collagen, muscle, and fat tissue, blood vessels, and nerves. Leather is made from the leather skin layer from which hairs, fat tissue, nonfibrous protein, and water have been removed and which has been stabilized by various process steps. In a humid environment, non-stabilized hides decay quickly and lose their malleability upon drying.

Leather treatment and enzymes. Immediately after their removal from the carcass, hides are conserved by removing water (e. g., by salting) to prevent microbial attack. In a so-called water shop, several steps ensue: during soaking, blood, dirt, salts, fats, and nonfibrous protein are removed by the addition of water, surfactants, and reducing agents. In this step, the leather skin reabsorbs water. Proteolytic enzymes support this softening process significantly: they help to remove pigments, fat, and sweat glands – a prerequisite to arrive at high-value "aniline leather" with a clear scar pattern. The enzymes used in this process must exhibit high proteolytic activity without attacking collagen. Trypsin (pancreatic extracts) and proteases from *Bacillus subtilis* (→20) or *Aspergillus sojae* (→16) are well suited for this purpose. In the following liming process, epidermis and remaining hairs are removed, and the skin is opened up for the later tanning procedure. Lime and sodium sulfite are especially suited for this step. In the traditional craft, they were applied by hand on the inside of the hide. Today, strongly alkaline reagents composed of calcium hydroxide and sodium sulfite are used, and alkali-stable proteases are added (e. g., from *Bacillus alkalophilus*). During the ensuing bating process, alkali is removed by the addition of ammonium salts or organic acids. Pancreatic enzymes, as well as neutral and alkaline bacterial or mold proteases, may assist in this step to remove remaining non-collagen proteins and to loosen the collagen for dying. Mold proteases are also used for loosening chromium-treated leathers ("wet blues"). Attempts to combine soaking, liming, and tanning into one processing step based on intense enzyme treatment have not yet been accepted in the industry. However, the use of enzymes for the duration of each step, has reduced water consumption by 50 %.

Enzymes for textile treatment. Weaving textile fibers exposes the fabric's warp and weft threads to substantial mechanical stress. To prevent breakage of the threads during this step, they are usually reinforced with sizing. For cost reasons, starch is usually used as the sizing agent. Polymer dispersions such as ethylen-vinylacetate or polyurethanes are also being used. Sizings adhere to the threads very well but must be removed after weaving, before the ensuing steps of dying, bleaching, and fortifying. In the case of starch, Bacillus α-amylase (→176) is usually used for this purpose. "Stone washing" is used to give denim textiles such as jeans a worn-out look. In the original method, textiles were washed with pumice stones. These have a rough surface and scrap off a layer on the denim so that some of the white threads from the cloth becomes more visible. As this procedure is quite aggressive to the textile, Aspergillus cellulases are now prefered. They remove some layers of dyed fiber from the surface without harming the textile. Cellulases such as BioBlast® are also being used for the "depilation" of cellulose fabrics or garments getting rid of protruding pills (→174).

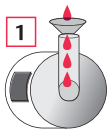
Leather manufacturing



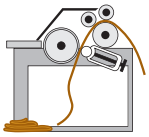
salt curing (preservation)

raw hide stock depot

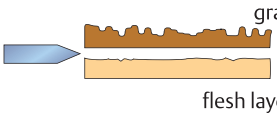
soaking, liming (hair loosening, opening up corium)



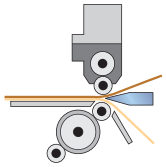
fleshing: hides in this state called "white" or "limed" stock



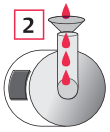
splitting skins: corium is cut horizontally by guiding skins along band knife splitter



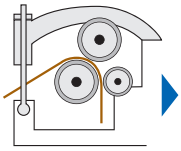
splitting skins



deliming, bating, pickling, chrome-tanning

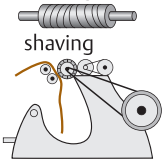


wringing



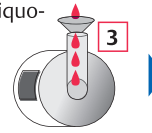
shaving knife

shaving

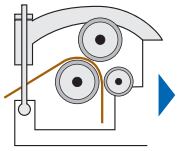


at this stage of manufacturing, leather is called "wet-blue"

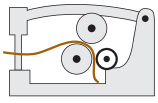
neutralizing, retanning, dyeing, fat liquoring



wringing

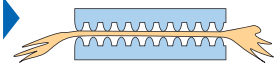


setting

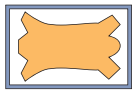


drying

conditioning (staking)



stretching

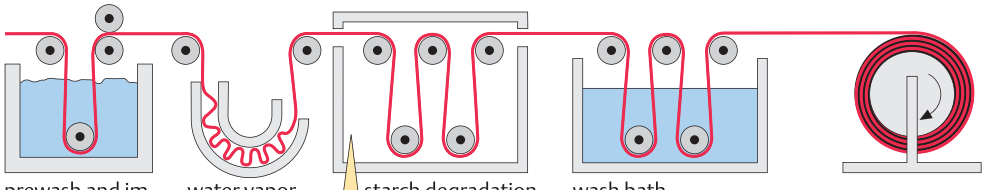


at this stage of manufacturing, leather is called "crust"

Enzymes used in leather manufacturing

	processing step	pH	target proteins	enzymes
1	soaking (degradation of proteins, opening up of corium, hair loosening)	7 – 9	non-fibrillary proteins, fats	proteases from molds, bacteria; trypsin (pancreas extracts)
1	dehairing	10	keratin	keratinases, proteases
1	liming	12.5	non-collagen proteins	proteases from bacteria, molds; elastase
2	deliming, bating	8 – 9	residual proteins, fat cells	proteases from bacteria, molds; trypsin (pancreas extracts)
2	pickling (enhancement of elasticity, improved dyeing)	5 – 6	collagen	proteases from bacteria, molds; trypsin (pancreas extracts)
3	wet blue	6	chromium-tanned leather	mold proteases

Desizing of textiles



prewash and impregnation with amylase

water vapor

starch degradation

wash bath

normal or thermophilic α-amylase from Bacillus strains, T = 60 –110 °C, depending on enzyme

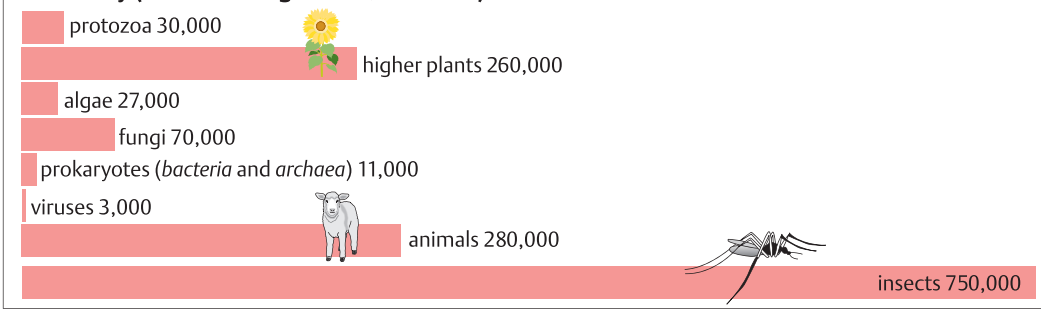
Procedures for obtaining novel technical enzymes

General. The isolation of new technical enzymes has accelerated, due to improved screening procedures, expression in host organisms that can be handled easily, and enhanced productivity using mutagenesis and selection or genetic engineering techniques. In spite of this progress, enzyme processes are only gradually overtaking procedures so far done by organic synthesis. One reason is that, for some important synthetic unit operations such as C-C coupling, only a few enzymes are available that do not require expensive cofactors; aldolases are one positive example. On the other hand, the time required to develop and optimize an enzyme process is often too long to suit the short time requirements encountered in the establishment of GMP-attested protocols as required in the pharmaceutical and agricultural industries (→334). Current ideas for improving this situation include 1) rational protein design (→198), 2) directed evolution (→198), and 3) novel procedures that are better suited to enhance the high natural diversity of biocatalysts, e. g., metagenomics (→74).

Novel habitats. Organisms are adapted to very diverse habitats (ecological niches) and usually have evolved enzymes that are perfectly suited to this environment. Microorganisms are especially versatile at adapting to habitats exhibiting a wide range of pH values, temperatures, and osmotic and other conditions. A systematic screening in unusual habitats such as highly acid or alkaline environments, salt deserts, deep-sea sediments, hot springs on the surface of the earth or on the bottoms of the oceans have allowed the identification of a large number of unusual microorganisms and the isolation of novel enzymes from them. For example, the thermostable DNAse I used in most PCR reactions (Taq polymerase) was found in the thermophilic prokaryote *Thermus aquaticus*, which was first isolated from the 90 °C waters of Old Faithful, a geyser in California's Yellowstone Park. Deep-sea submarine expeditions to hydrothermal vents 1500 m deep on the bottom of the Mediterranean Sea have led to the isolation of hitherto unknown microorganisms that belong to the group Achaea. Due to the extreme conditions in their environment (90 °C and 150 bar), their DNA polymerases show even better fidelity in DNA replication than that of *Thermus*. Some of these novel enzymes have become commercially available (e. g., Pfu- and Tma-polymerase) (→50).

Novel screening methods. Besides traditional screening methods, the modification of enzymes by protein design or directed evolution (→198) has become a powerful tool for optimizing an enzyme for its intended technical application. Directed evolution in particular has allowed rapid adaptation of enzymes to the desired temperature, pH, organic solvent or substrate. The flood of sequence information resulting from genome sequencing, analyzed by the tools of bioinformatics, has also opened up the potential to identify new enzymes. In the ever-increasing number of completely sequenced genomes, many thousands of enzymes could be identified due to their fingerprints (consensus sequences, as already known from enzymes with similar functions). They can be isolated rather easily using PCR techniques. The search for their natural substrate, however, can often be quite troublesome. By analyzing sequences in DNA samples isolated from soil which code for conserved regions of the 16S-, 18S- or 23S-ribosomal RNA, it has been concluded that <1 % of all microbial species have yet been cultivated and classified (→74). To circumvent this limitation, methods have been successfully established to isolate new enzymes from uncultivated microbes by isolating DNA from "metagenomes" in soil, sludges or rumen, cutting them with restriction enzymes, and expressing their genes or gene fragments in suitable host organisms, using appropriate functional assays or matching sequence information to the sequence of enzymes of known function. Bioinformatic procedures (→324) may also help to search for technical enzymes: if the desired substrate of an enzyme is known, substrate binding sites within the 3D-structures of the appropriate type of enzymes may be virtually docked with these substrates, leading to a ranking of the most probable enzyme candidates. The cloning and experimental validation of these virtual candidates is the logical next step. Many procedures for obtaining new enzymes are now offered by service providers such as B. R. A. I.N. AG, Evocatal, Enzymicals or BioProdict.

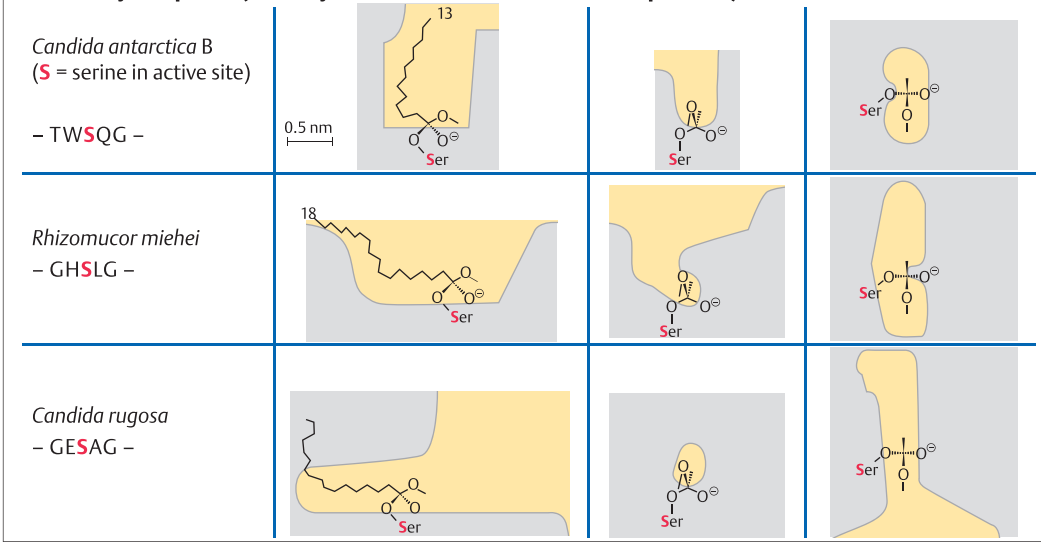
Diversity (all known organisms, estimate)



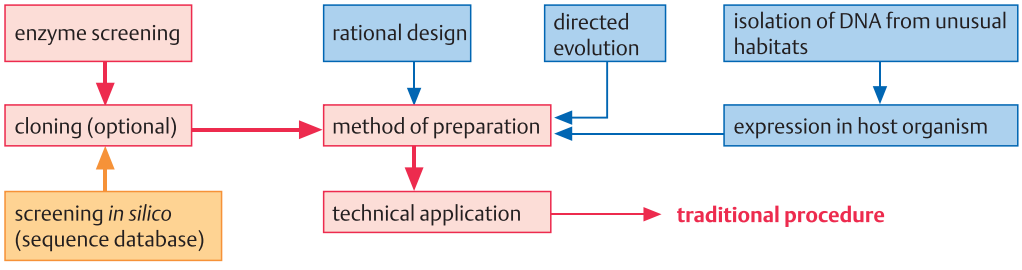
Properties of various DNA polymerases

origin	half life time at 95 °C (min)	3' → 5' - exo-nuclease activity	5' → 3' - exo-nuclease activity	type of terminus
Thermus aquaticus	40	not present	present	3'A
Thermotoga maritima	20	not present	present	3'A
Pyrococcus furiosus	> 120	present	not present	n.i.
Thermococcus litoralis	1 380	present	not present	> 95 % blunt end

Diversity of lipases (Conolly structures and consensus sequences)



Strategies towards new enzymes



DNA	restriction digest	expression in host organism	assay
from sediments, soil samples, sludge, etc.	into DNA fragments of the desired size	in *E. coli*, *Streptomyces*, and others	selection of hits

Protein design

General. Protein design, or protein engineering, implies the modification of protein sequences by genetic methods. Protein engineering techniques are used to 1) understand enzyme mechanisms, 2) modify enzyme or antibody binding sites at will, and 3) alter global properties of an enzyme, such as its stability to high temperature, extreme pH, proteases, its solubility, or its antigenicity. If a known protein structure is used as the starting point, and individual amino acids or sequences are replaced by site-directed mutagenesis, the protocol is termed rational protein design. The genetic exchange of amino acids at random and selection of hits by their improved properties is called directed evolution.

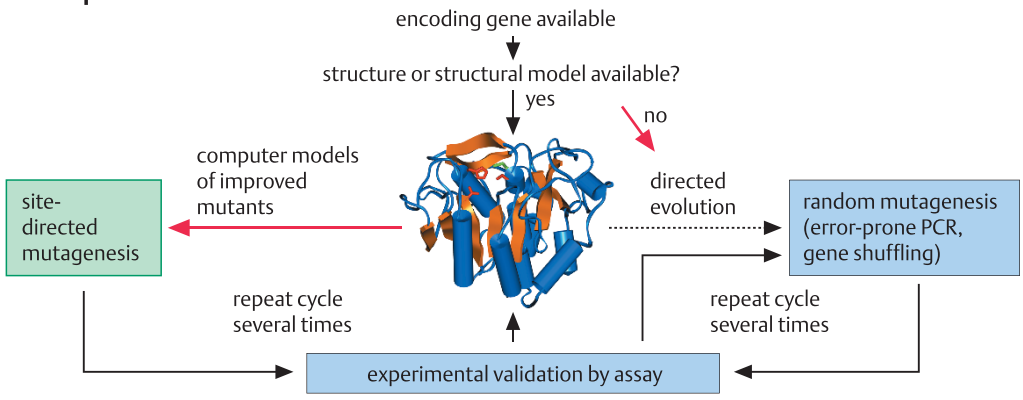
General methods. Both rational protein design and directed evolution require that the gene coding for the protein be available. For rational protein design, structural information about the protein is required; this can be derived from the x-ray structure (→28, 324), from NMR structural data, or from a structural model that was derived from the tertiary structures of closely related proteins by homology modeling.

Mutagenesis. In rational protein design, individual amino acids are exchanged, or an amino acid sequence is introduced or deleted. This is done at the level of DNA by PCR (→52). Several protocols are available that enable these modifications with rapid, simple, reliable procedures. For random mutagenesis, the gene can be cloned into an *E. coli* host strain whose DNA repair mechanisms have been impaired and which is cultured under mutagenic conditions. Alternatively, a PCR protocol is used in which addition of Mn^{2+} ions or other means leads to an artificially high number of errors during DNA amplification (1–3 %). Gene shuffling, yet another method, is based on the concept of creating a library of DNA fragments of related genes (their sequence identity must be ca. 80 %) and recombining the fragments by PCR methods, followed by high-throughput screening for the desired properties.

Rational protein design. The tertiary structure of a protein is usually obtained by x-ray crystallography, sometimes also by 3D NMR techniques from ^{13}C- and ^{15}N- labeled proteins. As of 2014, coordinates of >100,000 protein structures are available from ProteinDatabase (PDB), which are internationally accessible through the Internet. If a protein of interest shows sequence homology of >30 % to a protein whose coordinates are available, homology modeling of the unknown structure based on the known coordinates usually provides a structural model of the unknown protein that is sufficiently precise for mutagenesis experiments. Until recently, such simulations were only possible in vacuum, due to limited computer power. With the advent of supercomputers and highly parallel computers, the modeling of protein structures, of mutant proteins, and of their binding to substrates or antigens can be done in solvent (this may require molecular mechanics calculations (forcefield calculations) of the interactions of several tens of thousands of atoms!). Although improving, the predictions derived from such *in silico* methods must usually be optimized by several cycles of simulations and genetic experiments (protein engineering cycles). Protein design is often combined with a permutation of all proteinogenic amino acids in positions that have been chosen by substrate docking or by analogy considerations as relevant sites for, e. g., the binding of a substrate ("saturation mutagenesis").

Directed evolution. In contrast to rational protein design, structural models are not necessary for this technique. For optimizing the binding selectivities of antibodies, the phage display technique (→244) has been effectively used: it allows large libraries of mutant antibodies (up to 10^{10}) to be screened. In the case of enzymes, the encoding gene is subjected to random mutagenesis, the mutant genes are expressed in a mutant library, and the mutants are assayed for the desired properties. By iterative saturation mutagenesis programs, more targeted approaches for probing the huge protein sequence space have become rapidly available. The quality of the enzyme assay is of crucial importance and determines both the speed of obtaining and the quality of the mutants. In recent years, this method has yielded excellent results for the improvement of industrial enzymes, e. g., in altering their substrate specificity or their thermo- and alkali stability. Robotic systems (→310) or FACS equipment (flow-activated cell sorters) (→84) are generally used for these methods.

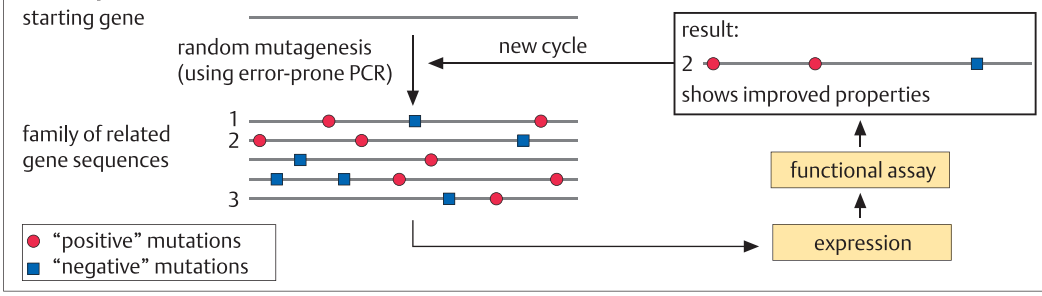
Concepts



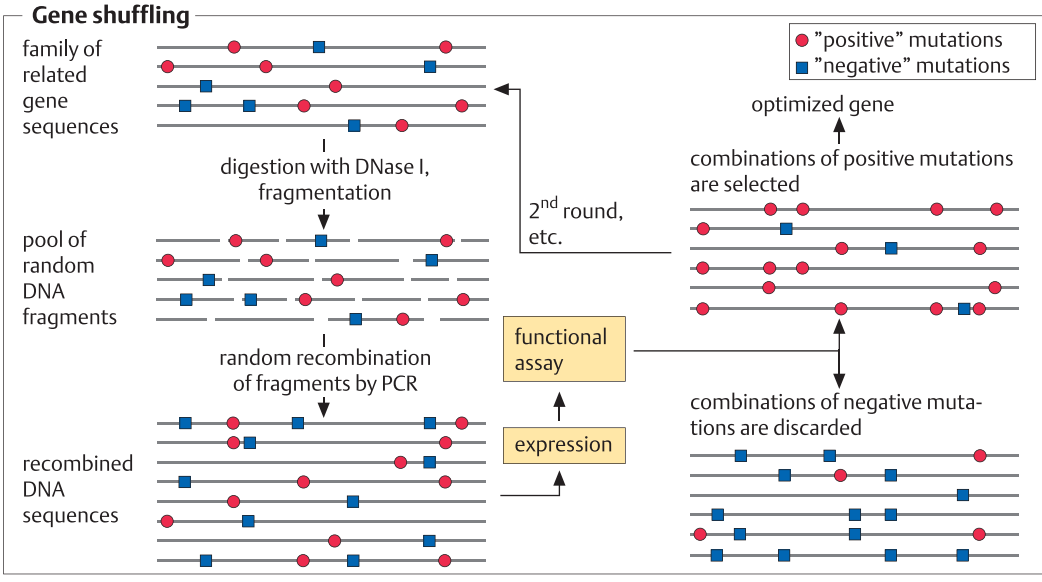
Examples

enzyme	method	mutation	effect
detergent protease	rational design	^{222}met→ala	oxidation stable
human insulin	rational design	^{22}lys→pro	slower degradation
tissue plasminogen activator (hTPA)	rational design	deletion of kringle 2	slower degradation
penicillin acylase	directed evolution	5 positions	better solvent stability
P450 fatty acid hydroxylase	directed evolution	4 positions	greatly modified substrate specificity

Error-prone PCR



Gene shuffling



Antibiotics: occurrence, applications, mechanism of action

General. In 1928, the British microbiologist Alexander Fleming noticed that a fungal infection on his agar plate inhibited the growth of Staphylococci. Only a decade later, Howard Florey, an Australian working in the UK, succeeded in the isolation, purification, and structural identification of the mixture of fungal antibiotics responsible for Fleming's observation. Successes in animal experiments and in treating a patient severely affected by a Streptococci infection led to the initiation of a large-scale project by the British/US allies in World War II. By 1945, penicillin could be prepared in kg amounts. In 1947, Selman Waksman, a Russian-American microbiologist, detected streptomycin, a new antibiotic in cultures of *Streptomyces griseus* that was effective against Gram-negative microorganisms in addition to penicillin. In the following years, systematic screening led to the discovery of very many new antibiotics, and processes for their manufacture in the industrial scale became established. However, the uncritical use of antibiotics against trivial diseases and as a feed component in mass animal production soon led to the development of antibiotic resistance in microorganisms, which is now a major concern in hospitals ("nosocomial infections") (→204). In an attempt to counteract this phenomenon, novel types of antibiotics are being developed and their use is being restricted to specific applications. Screening methods have recently been complemented by genetic methods such as cell fusion of different antibiotic producers, combinatorial biosynthesis and shuffling of gene clusters responsible for the biosynthesis of antibiotics. Further, chemical modification of known antibiotics that were not used yet due to drawbacks in toxicity or productivity have improved development.

Antibiotics as bioactive molecules. Antibiotics are secondary metabolites (→36), a large group of molecules that are formed from the building blocks of central metabolism but are extensively modified after growth has stopped. Their average molecular weight is 400–600 Da. Major biological roles of secondary metabolites are in defense, regulation and communication. Their functions may include enzyme inhibition, pheromone activity, petal coloring or toxin activity. So far, >500,000 secondary metabolites have been identified, most of them in higher plants. Antimicrobial activity of secondary metabolites may be just one aspect of their function; in fact, over half of all antibiotics show other biological effects as well. The genome sequences of many microorganisms that produce antibiotics have been solved, providing clues to the regulation of antibiotic synthesis and self-protection. Usually, both biosynthesis and resistance are coded by genes that are combined in one or a few operons.

Occurrence. ~10,000 antibiotics have been isolated from microorganisms, another 10,000 from basidiomycetes and lower fungi, and 10,000 more from higher organisms, especially plants but also lichen and animal and marine invertebrates. Actinomycetes by far outnumber all other organisms in their capacity to synthesize antibiotics.

Applications. Only ~300 antibiotics are produced industrially. They are mostly semisynthetic compounds, in which a biologically active lead structure is modified chemically. β-Lactam antibiotics (penicillins and cephalosporins) (→206) make up about half of a world market of >42 billion US$ and some 50,000 t (2009). Most antibiotics are manufactured as antimicrobial agents for chemotherapy. They can be classified as broad-spectrum antibiotics, affecting a wide range of pathogens (e. g., cephalosporins, tetracyclines), and selective antibiotics, used for highly special therapies (e. g., rifampicin against tuberculosis, amphotericin B against fungal infections). Antitumor antibiotics, such as adriamycin, are valuable cytostatic agents, but also exhibit high toxicity. Several antibiotics are used for plant protection. They are often effective in lower concentrations than herbicides and show low toxicity against mammals. Examples are blasticidin S and kasugamycin. Only a few antibiotics are used in food preservation, such as pimaricin, which is sometimes used as an antifungal agent in cheeses. Feed antibiotics lead to a better mast performance and faster growth in mass animal production, but are usually restricted for feed use (e. g., monensin in broiler production), to prevent clinical cross resistance. In molecular biology, antibiotics serve as a research tool for the selective inhibition of various cell functions.

Occurrence

taxonomic group	relative number (%)*
Actinomycetes	40
other bacteria	15
fungi	20
algae	1
plants	12
marine invertetebrates	10
animals	1

* ~30 000 active substances (estimate)

Production and value

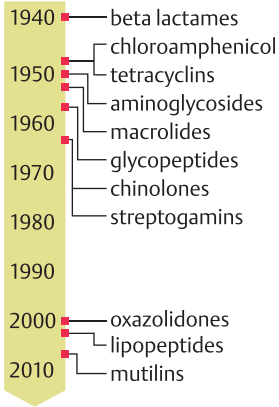
type of antibiotic	value [10^9 US$]
glycopeptides, lipopeptides	0.2
aminoglycosides	1.0
tetracyclins	1.6
macrolides	4.8
fluorochinolones (synthetic)	7.1
penicillins/lactamase inhibitors	11.5
cephalosporins	11.9
other	3.9
total	42

(2009)

Classification by chemical structure

1	**carbohydrate antibiotics**	aminoglycosides	streptomycin (medicine), kasugamycin (rice fungicide)
2	**macrocyclic lactones**	macrolides polyene antibiotics ansamycines	erythromycin (medicine) pimaricin (cheese production) rifamycin (against tuberculosis)
3	**chinones and related antibiotics**	tetracyclines anthracyclines	tetracyclin, chlorotetracycline (medicine, feed antibiotic) doxorubicin (cancer therapy)
4	**amino acid and peptide antibiotics**	amino acid derivatives β-lactam antibiotics peptide antibiotics chromopeptides glycopeptides	cycloserine (organ transplantation) phosphinothricin (plant protection) penicillins, cephalosporins (medicine) bacitracin (medicine), virginiamycin (feed antibiotic) actinomycin (cancer therapy), bleomycin (cancer therapy), vancomycin (medicine), avoparcine (cattle feed antibiotic)
5	**N-heterocyclic compounds**	nucleoside antibiotics	polyoxins, blasticidin S (fungicides for plant protection)
6	**O-heterocyclic compounds**	polyether antibiotics	monensin (chicken feed)
7	**alicyclic compounds**	cycloalkane derivatives	cycloheximide (leaf fungicide)
8	**aromatic antibiotics**	benzene derivatives	chloramphenicol (medicine) griseofulvin (fungicide)

Time of introduction



4	amino acid and peptide antibiotics	lipopeptides	ciclosporin β (transplantation medicine) daptomycin (therapy)
5,6	N- and O-heterocyclic compounds	oxazolidinone	linezolid (therapy)
7	alicyclic compounds	mutilins	retapamulin (therapy)
8	benzene derivatives		

New antibiotics by
- screening for new lead structures
- modification of old and new lead structures
- modification of already known but yet unused antibiotics

Antibiotics: screening, industrial production, and mechanism of action

Screening. Traditionally, inhibition of microorganisms' growth in the presence of a producer or its culture filtrate is used to test for an antibiotic. If an interesting biological activity is detected, the antibiotic is enriched and purified from the culture broth, and its structure is elucidated. If this procedure is used today, an already known antibiotic is usually rediscovered. Thus, to increase the number of "hits," many new screening procedures have been developed. For example, other biochemical or biological assays may replace growth inhibition tests using microorganisms, culture filtrates of putative antibiotic-forming strains may be analyzed by chemical procedures, or target-oriented methods may be used (e. g., the inhibition of a microbial enzyme involved in pathogenesis is monitored in a highly parallel microtiter plate assay against whole or fractionated extracts of producer strains).

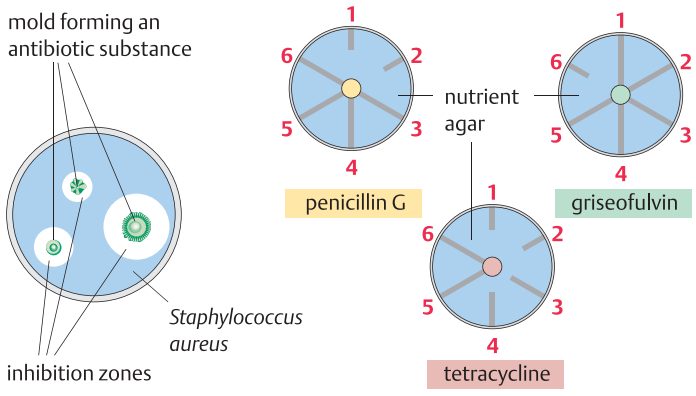
Strain improvement. If a promising new antibiotic is identified, its yield must be optimized at an early stage, since antibiotics are secondary metabolites and thus formed in rather low concentrations (several mg L^{-1} culture or less). Strain improvement usually follows empirical rules dominated by labor-intense repetitions of mutation and selection. Occasional back-crossing with the wild-type strain may enhance the robustness of production strains. Based on such methods, yields of economically important antibiotics were increased 10^3–10^6 fold relative to the original isolate. Genetic engineering techniques, e. g., increasing the number of gene copies (→62) for key enzymes, have led to further improvements in yield.

Fermentation and recovery. The chemical constitution of most antibiotics is complex. They usually contain several stereocenters, rendering chemical synthesis complicated. In fact, the syntheses of many antibiotics have been brilliant examples of chemical synthesis that are not applicable to industrial production. As a result, most antibiotics are produced by fermentation in bioreactors. Inexpensive carbon and nitrogen sources such as molasses, dextrose, whey, soybean meal, and corn steep liquor are used. Since most antibiotic producers are subject to catabolite repression, the carbon source (→88) is often added in a fed-batch mode. Because antibiotics are secondary metabolites, their formation starts only after cell growth has reached its stationary phase (→90). Antibiotics are usually extracellular products, and often they are only moderately soluble in water. For their isolation at the end of the fermentation process, cells are removed and the culture broth is extracted with organic solvents (the two steps can be integrated into one) (→104). The raw antibiotic can be further purified by various chromatographic procedures or by crystallization. The manufacture of antibiotics is subject to strict regulations and safety controls following certified rules, e. g., GMP and ISO 9000.

Mechanism of action. Antibiotics can affect: 1) biosynthesis of the components and function of the genetic machinery of the cell, 2) biosynthesis of cell components, 3) biosynthesis and function of proteins, 4) biosynthesis and function of the cytoplasmic membrane or, as in Gram-negative bacteria, the outer cell membrane, and 5) biosynthesis of the cell wall. The interactions of individual antibiotics are diverse. In view of toxic side effects, those antibiotics that act on targets specific to microorganisms are preferred as drugs, such as the lactam antibiotics that inhibit the biosynthesis of murein, the peptidoglycan polymer of the bacterial cell wall. As a consequence of their short generation time and their rapid adaptation to changing environments, microorganisms usually become quickly resistant to antibiotics, resulting in a permanent race between the development of new antibiotics and the occurrence of new resistant strains.

Target-Screening. The genomes of well over 2000 microorganisms have now been sequenced (2014), among them many pathogens. Such greatly enhanced knowledge about the components of microbial cells, their metabolism and regulation has raised hopes to identify more specific targets for the development of new synthetic molecules with antimicrobial activity. Thus, in a seven years program, 300 potential targets were prepared, and millions of drug candidates were tested on them using high-throughput assays. The results were rather disappointing as only five putative new lead structures were identified. As a consequence, the present focus is rather on the chemical modification of already known antibiotics or on the generation of new lead structures using combinatorial biosynthesis.

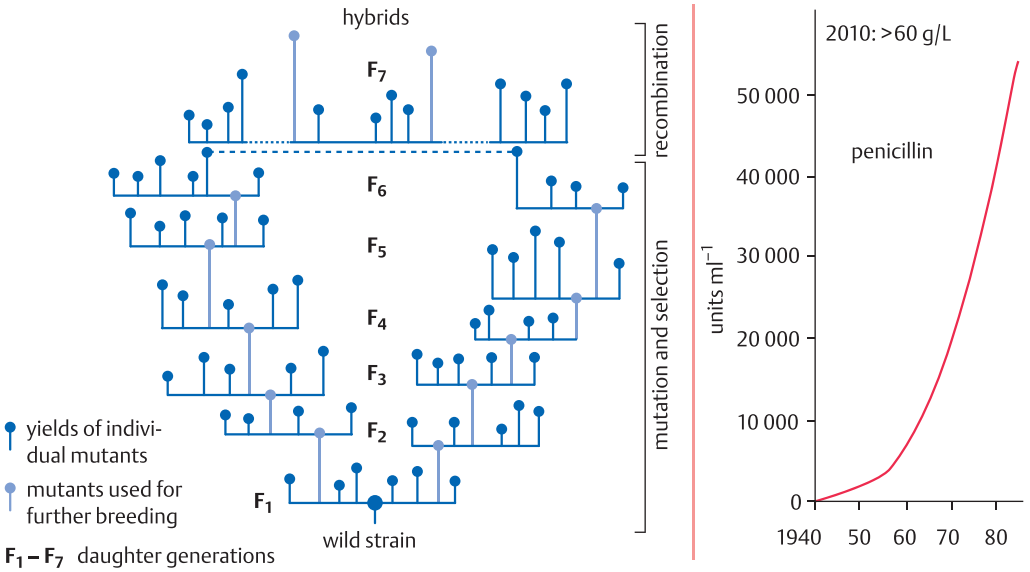
Screening



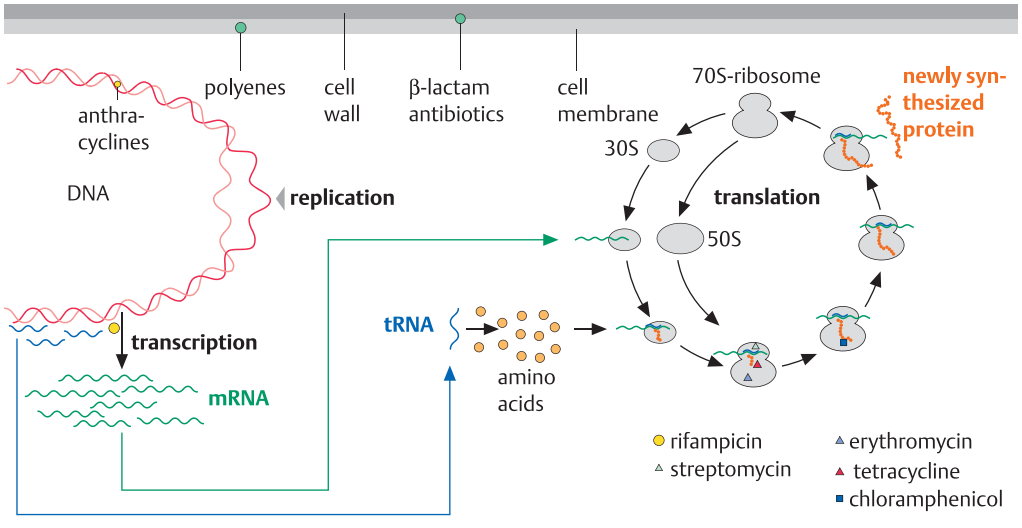
test organisms:

1. *Staphylococcus aureus* (Gram-positive bacterium)
2. *Streptococcus* sp. (Gram-positive bacterium)
3. *Escherichia coli* (Gram-negative bacterium)
4. *Pseudomonas aeruginosa* (Gram-negative bacterium)
5. *Candida albicans* (yeast)
6. *Trichophyton rubrum* (fungus)

Strain improvement



Targets of antibiotics



Antibiotic resistance

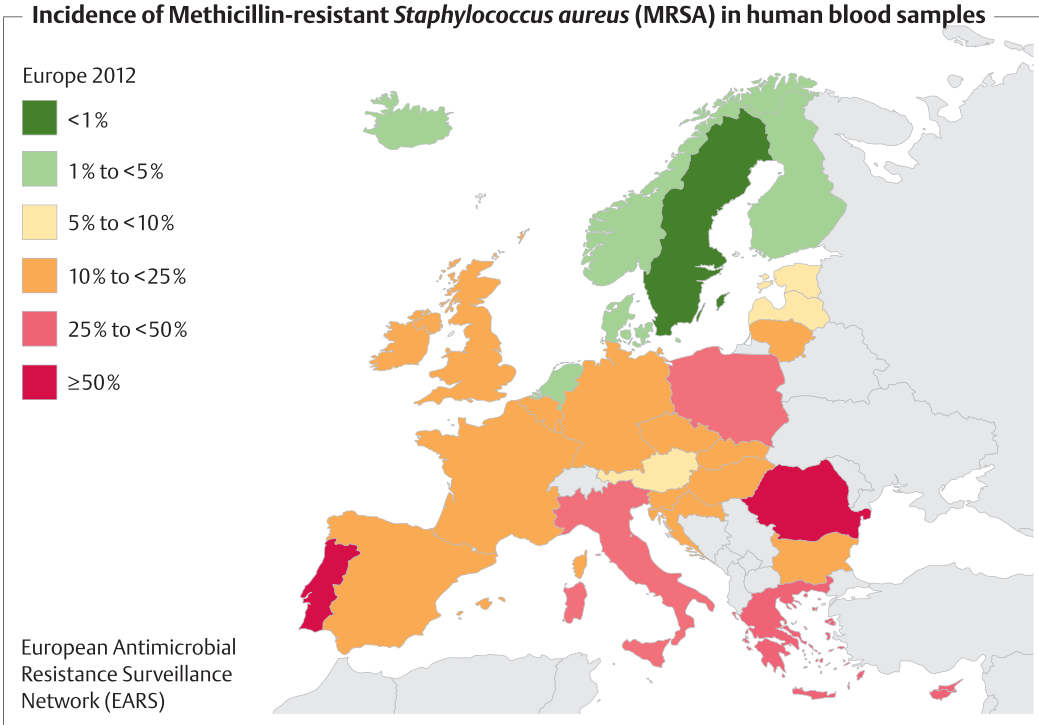
Resistance. The rapid spread of microorganisms that are resistant to antibiotics is a major medical problem. Multi-drug resistant strains (i. e., resistant against many antibiotics) of Salmonella, *Escherichia coli*, Streptococci and Staphylococci have been widely propagated. The number of humans infected with a latent form of *Mycobacterium tuberculosis*, the causative agent of tuberculosis, is estimated at close to 2 billion people. Today, over 70 % of all human pathogens carry resistance against most antibiotics under clinical use, and the global number of victims is estimated at 2 million people per year (2013). If patients are infected with Methicillin-resistant *Staphylococcus aureus* strains (MRSA) or Carbapenem-resistant Enterobacteria (CRE), mortality is over 50 %. As a result, global surveillance systems were installed which follow up on the spread of resistant microorganisms. In the United States this system is the National Antimicrobial Resistance Monitoring System (NARMS) of the FDA; in Europe it is the European Antibiotic Resistance Surveillance System (EARSS). Resistance mechanisms are genetically encoded, and genes conferring resistance can be transferred among different microorganisms by plasmids, phages or transposons. As a consequence, there is an endless race to develop novel antibiotics to which microorganisms do not (yet) show resistance. In addition, there are strict regulations on those antibiotics that are used in industrial livestock farming as growth promoters or for prophylaxis. They are now banned in those cases where it has been proven or is highly probable that they move through the food chain and cause multidrug resistance in human isolates.

Mechanisms of resistance. The most frequent mechanisms of resistance are a) reduced resorption or fast export of the antibiotic (e. g., by modulation of membrane permeability or exporter pumps), b) modification of the targets (e. g., through variation of the antibiotic binding site at the ribosome or the DNA), and c) enzymatic modification of the antibiotic. At the level of microbial communities as they occur, e. g., in biofilms, in the gastrointestinal tract of man and animals or in sewage plants, an additional mechanism lies in *quorum sensing* which initiates, through chemical messenger compounds such as acyl-homoserine lactone (AHL), the genetic onset of resistance formation. Antibiotics that attack single targets (e. g., ribosomes by macrolides) induce resistance faster when compared to those antibiotics that attack broader targets (e. g., lactam antibiotics). When new antibiotics are developed, it is attempted to prevent or at least slow down resistance formation. To this end, chemical modifications of a basic antibiotic structure *via* chemical synthesis, precursor-directed or combinatorial biosynthesis is used (→220).

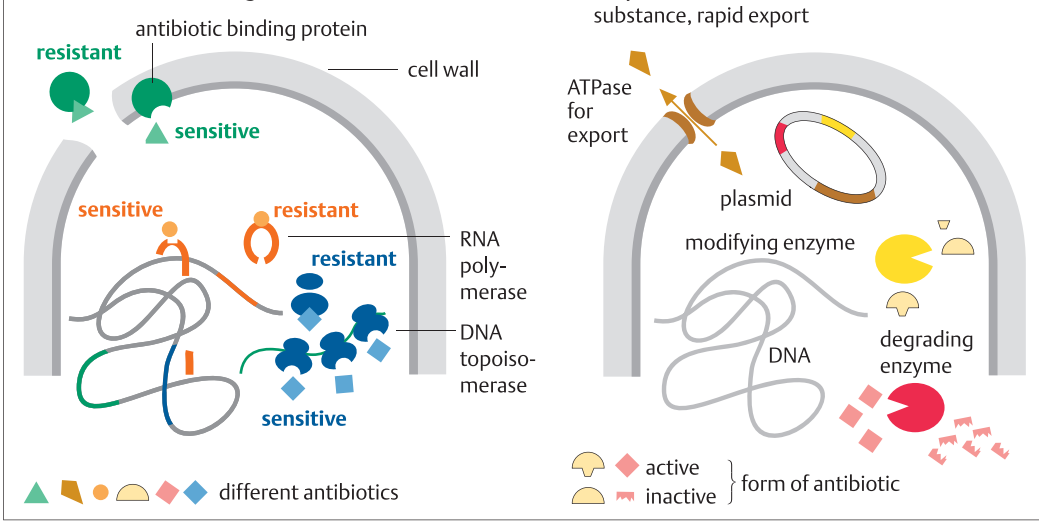
Clinical aspects. Hospital infections ("nosocomial infections") are frequent and dreaded events. Preventive measures include rigorous hygiene, medical guidelines in the selection of antibiotics, and patient screening upon hospitalization using rapid diagnostic tests. Among the latter, rapid tests for CRE, MRSA, *M. tuberculosis*, *Clostridium difficile*, Rifampicin-resistant strains and infectious viruses (norovirus, adenovirus) dominate.

Diagnostic procedures. The diagnosis of an antibiotic resistant microorganism is traditionally done by cultivation of a clinical sample (e. g., sputum) in a microbiological laboratory. In a first step, the suspected strain is isolated on a selective nutrient agar, to be tested in a second step, e. g., by a disc plate assay, against various antibiotics (→202). In this method, paper strips each loaded with a different antibiotic are put on a growth agar, and growth inhibition zones are observed. In spite of various automation procedures, this method requires at least 2 days before results are obtained, which may be critical in the case of acute life-threatening infections and does not allow for the prevention of nosocomial infections through the rapid diagnosis of new patients to a hospital. Significantly faster methods are becoming available using RT-PCR on the DNA of a suspected microorganism (→302). Fully automatic systems are already on the market which allow the isolation of sample DNA from a sputum probe, amplification of the relevant DNA-segments using appropriate primers, and quantitative determination of the amplicons to provide results in less than two hours. More complex assays for the rapid genotyping of, e. g., lactam antibiotic resistance are under investigation.

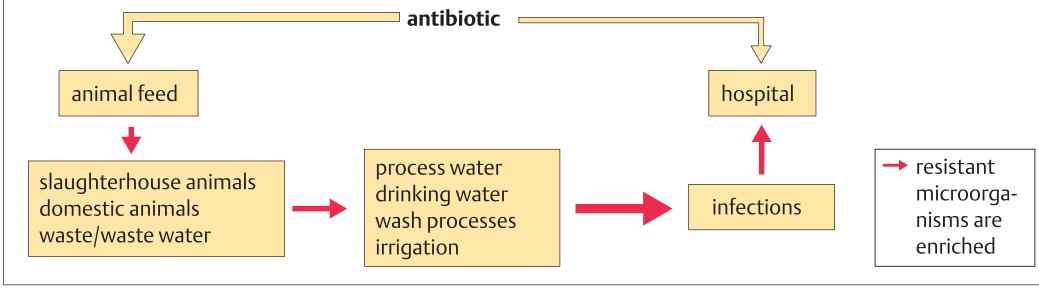
Incidence of Methicillin-resistant *Staphylococcus aureus* (MRSA) in human blood samples



Mechanisms of resistance



Antibiotics as growth promotors in animal feed: the problem of resistance



β-Lactam antibiotics: structure, biosynthesis, and mechanism of action

General. β-Lactam antibiotics (*Penams*: penicillins. *Cephems:* cephalosporins) are the most important antibiotics both in value and quantity, due to their high efficiency, low toxicity, and wide variety of ways to turn them into semi-synthetic lactam derivatives. Annual production is in the range of 6 x 10^4 tons; cephalosporins constitute about 50 % of the total antibiotic market. They are exclusively used in clinical applications. In contrast, penicillins are also used in veterinary medicine. The most important primary β-lactams are penicillin G and cephalosporin C. They serve as starting materials for the manufacture of semi-synthetic penicillins and cephalosporins. Important criteria for the efficacy of β-lactam antibiotics include acid stability (for oral delivery followed by stomach passage) and stability against the plasmid-encoded β-lactamases, that are a key factor responsible for microbial resistance to β-lactam antibiotics.

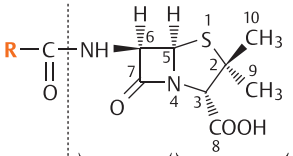
Penicillins. *Penicillium chrysogenum* (→16) forms isopenicillin N as the primary product. It contains a nonproteinogenic amino acid in its side chain (L-α-aminoadipic acid). If acids, e. g., phenylacetic acid, are added to the culture medium in late logarithmic stage, a fungal N-transacetylase catalyzes the incorporation of these acids into "biosynthetic penicillins" that exhibit different pharmacological properties. Thus, if phenoxyacetic acid is added to a *P. chrysogenum* culture in late logarithmic phase, Penicillin V is produced. Penicillin V is acid-resistant and can be orally administered. Penicillin G and 6-aminopenicillanic acid (6-APA), the intermediate obtained by hydrolysis of the acid side chain, are the most important intermediates for making semisynthetic penicillins.

Cephalosporins. Following observations by Giuseppe Brotzu, an Italian physician, cephalosporin C 1953 was the first to be discovered as a new class of β-lactam antibiotics. It is formed by *Acremonium chrysogenum* (former name: *Cephalosporium acremonium*). *A. chrysogenum* does not contain an N-transacetylase; in consequence, it is impossible to prepare biosynthetic cephalosporin derivatives from added acid precursors. Most semisynthetic cephalosporins are therefore produced from the synthetic intermediates 7-amino cephalosporanic acid (7-ACA) and its deacetylated derivative 7-amino desacetoxy cephalosporanic acid (7-ADCA). Cephalosporins of the 2nd and 3rd generations show an excellent, broad efficacy towards a wide range of Gram-positive and Gram-negative pathogens at low human toxicity. ß-Lactamase inhibitors such as clavulanic acid are sometimes used in conjunction with lactam antibiotics to expand activity.

Biosynthesis. *P. chrysogenum* contains a cluster of three genes that code for the biosynthesis of isopenicillin N. The building blocks L-α-aminoadipic acid, L-valine, and L-cysteine are condensed by a single synthase to a tripeptide, which is transformed by a second synthase into bicyclic isopenicillin N. An acyl-CoA:isopenicillin-N-acyl transferase coded by the same gene cluster allows for the exchange of the L-amino acid side chain against a wide range of other acyl groups. The biosynthetic pathway is distributed among several cell compartments. Cephalosporins are formed in *A. chrysogenum* from intermediary isopenicillin N, after epimerization of the amino acid side chain, via an oxidative ring expansion catalyzed by expandase, followed by reactions on various substituents of the lactam and the thiazine ring. The O-acyltransferase of the fungus allows modification reactions at the 3-acetoxymethyl position, but the lack of an N-acyltransferase prevents reactions at the 7-amide bond. In contrast to Penicillium, the genes for the biosynthesis of cephalosporin C are located on two chromosomes and their transcription is strongly regulated. All the genes mentioned above have been cloned and their product formation is being experimentally modified via genetic and protein engineering.

Mechanism of action. β-Lactam antibiotics prevent the formation of peptide crosslinks during biosynthesis of the bacterial cell wall (murein) (→202). Since murein is a major component of the cell wall of Gram-positive microorganisms, it seems clear why the first β-lactam antibiotics where specific towards this group of bacteria. Higher animals and humans do not contain mureins, thus the side effects of β-lactam antibiotics are limited to interference with gastrointestinal microorganisms and to the occasional development of an allergy.

Penicillins

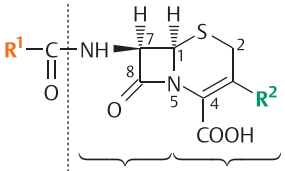


acyl residue — β-lactam ring — thiazolidine ring

6β-aminopenicillanic acid (6APA)

R	name	properties
HOOC—CH(NH$_2$)—(CH$_2$)$_3$— (S)	isopenicillin N	
H$_5$C$_6$—CH$_2$—	penicillin G	acid labile, sensitive to β-lactamase
H$_5$C$_6$—CH(NH$_2$)— (R)	Ampicillin	acid stable, sensitive to β-lactamase, active against gram-negative pathogens
HO—C$_6$H$_4$—CH(NH$_2$)— (R)	Amoxicillin	acid stable, sensitive to β-lactamase, broad-spectrum antibiotic, very well resorbed

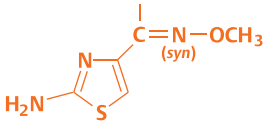
Cephalosporins



acyl residue — β-lactam ring — dihydrothiazine ring

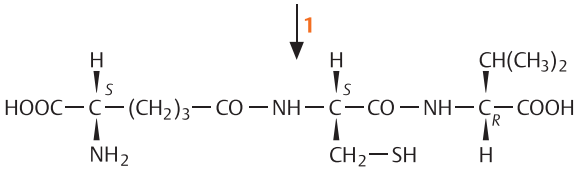
7β-amino cephalosporanic acid (7-ACA, R^2 = CH$_2$—O—CO—CH$_3$)

7β-amino desacetoxy-cephalosporanic acid (7-ADCA, R^2 = CH$_3$)

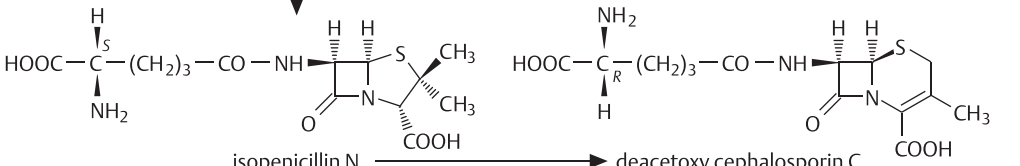
R^1/R^2	name	properties
HOOC—CH(NH$_2$)—(CH$_2$)$_3$— (R) CH$_2$—O—CO—CH$_3$	cephalosporin C	acid labile, sensitive to β-lactamase
H$_5$C$_6$—CH(NH$_2$)— (R) Cl	Cefaclor	acid stable, stable against β-lactamase, broad-spectrum antibiotic
 CH$_2$—O—CO—CH$_3$	Cefotaxim	acid stable, stable against β-lactamase, very broad efficacy

Biosynthesis and gene structure

L-α-amino adipic acid + L-cysteine + L-valine



δ-L-α-amino adipyl-L-cysteinyl-D-valine



isopenicillin N → (4, 5) deacetoxy cephalosporin C

3 → penicillin N

6 → cephalosporin C

enzyme/gene			
1	ACVS	*pcbAB*	δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine synthetase
2	IPNA AT	*pcbC*	isopenicillin-N synthetase
3		*penDE*	acyl-CoA: isopenicillin-N-acyl transferase
4	IPNE REX/H	*cefD*	isopenicillin-N epimerase
5	AT	*cefEF*	ring-expandase/hydroxylase
6		*cefG*	acyl-CoA: cephalosporanic acid acyl transferase

gene structure

P. chrysogenum

pcbAB *pcbC* *penDE*

A. chrysogenum

pcbAB *pcbC* *cefD* *cefEF* *cefG*

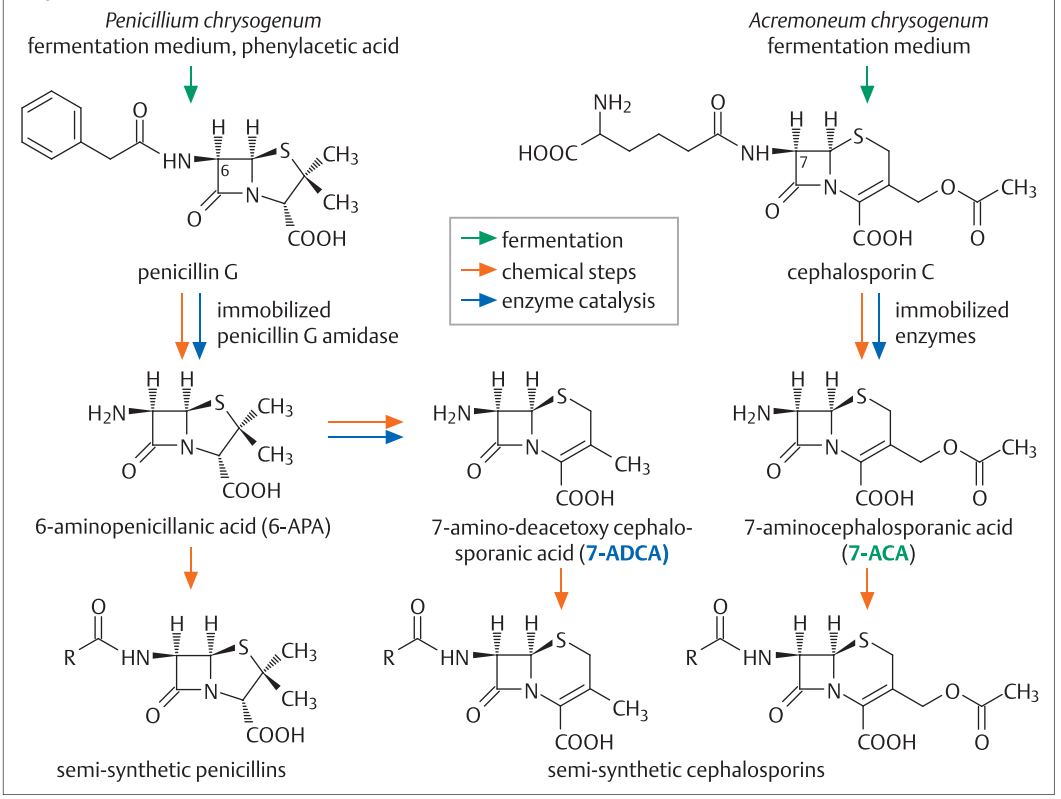
β-Lactam antibiotics: manufacture.

Manufacture of penicillins. The penam ring of the penicillins contains three stereocenters, and only one of the nine possible isomers [3(*S*):5(*R*):6(*R*)] is biologically active. As a result, a fermentation process is economically advantageous, relative to chemical synthesis. ß-Lactam antibiotics are formed by a number of bacteria and fungi, but only high-performance strains of *Penicillium chrysogenum* (→16) are under practical use. A range of aromatic or aliphatic precursors may be added to the culture medium, leading to the formation of "biosynthetic penicillins;" only penicillins G and V are produced in large scale, using phenylacetic acid or phenoxyacetic acid, respectively, as the precursors. Both penicillin G and V are also used in the manufacture of 6-aminopenicillanic acid (6-APA), the key intermediate for preparation of semisynthetic penicillins and some cephalosporins.

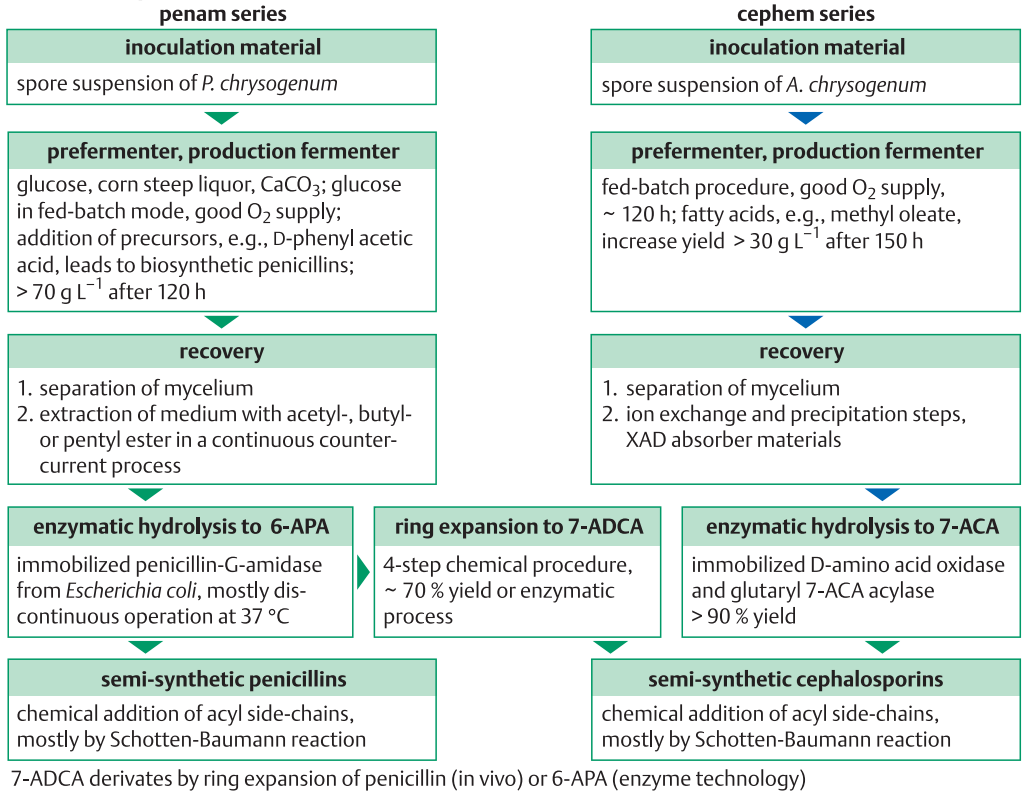
Penicillin G and 6-APA. For the industrial manufacture of penicillin G, high-performance strains of *P. chrysogenum* are grown in bioreactors of up to 200 m^3 volume (→92). To prevent catabolite repression, sugar is added via a fed-batch fermentation process. The oxygen supply (→94) is critical and requires careful optimization of the stirrer and the aeration system, since the fungal mycelium is both viscous and sensitive to shear. The complex nutrient broth contains glucose or saccharose as a C-source (→88) and corn steep liquor as an N-source. Formation of penicillin occurs mainly after ~40 h, when fungal growth has been completed. During the production phase of ~100 h, phenylacetic acid is fed to the medium, and penicillin G is secreted by the fungus. When fermentation is complete, the mycelium can be separated by filtration or centrifugation, and the filtrate is extracted within minutes with amyl or butyl acetate at 0–3 °C and pH 2.5–3.0 using a two-stage countercurrent extractor (→104). Alternatively, both steps can be combined by using a direct extraction procedure with two extraction decanters working in countercurrent mode. After re-extraction, using aqueous ammonia or carbonate, the crude antibiotic (> 3 tons for a 110 m^3 bioreactor) is purified by crystallization. Subsequent gentle hydrolysis of the amide bond leads to 6-APA, and phenylacetic acid that can be reused for fermentation. 6-APA is a key intermediate for the preparation of most semisynthetic penicillins and some cephalosporins. Over the past few decades, enzymatic hydrolysis of penicillin G, using immobilized penicillin G amidase from *E. coli*, has largely replaced chemical hydrolysis (→102). It is carried out at ~30 °C and pH 7.5–8.0, usually in batch mode. The high stability of the enzyme allows repetition of this step up to 1000 times before the enzyme must be replaced. Precipitation of 6-APA, filtration, and washing leads to a very pure product, which is further processed to semisynthetic penicillins or, through (chemical) ring expansion, to 7-ADCA, the base intermediate for some cephalosporins.

Cephalosporins and 7-A(D)CA. The fermentation procedure leading to cephalosporin C, using *Acremonium chrysogenum*, resembles that of penicillin G manufacture, but yields are lower. *A. chrysogenum* does not contain a N-transacetylase and thus cannot form biosynthetic cephalosporins. 7-Aminocephalosporanic acid (7-ACA) is prepared by hydrolysis of cephalosporin C, and the deacetylated analogue 7-ADCA by hydrolysis of glutaryl-3--desacetoxy cephalosporin. In the enzymatic procedure, which is more and more preferred for reasons of its favorable ecobalance, the D-α-aminoadipyl side chain is oxidatively deaminated into α-ketoadipyl-7-ACA by the action of an immobilized D-amino acid oxidase (→102). Spontaneous decarboxylation leads to glutaryl-7-ACA, from which the glutaryl side chain is cleaved off by an immobilized glutaryl-7--ACA-acylase. Cephalosporin C amidases, that can hydrolyze the D-α-aminoadipyl side chain in one step, have been reported, but their large-scale application is not yet established. A current option is to change the substrate specificity of glutary-7-aminocephalosporanic acid acylase into a cephalosporin C amidase. Genetically engineered strains of *P. chrysogenum* have been shown to produce cephalosporin C if an expandase/hydroxylase, cloned from *A. chrysogenum* or from *Streptomyces clavuligerus,* is expressed under the control of the β-tubulin promoter in the presence of adipic acid. Ring expansion of penicillin to desacetoxy cephalosporins has recently become possible by the use of the immobilized enzyme.

Synthetic routes



Process steps



7-ADCA derivates by ring expansion of penicillin (in vivo) or 6-APA (enzyme technology)

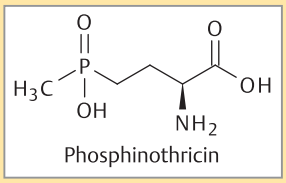
Amino acid and peptide antibiotics

General. β-Lactam antibiotics are the most important therapeutic antibiotics in human medicine and are discussed above. Among the >500 antibiotics originating from secondary amino acid metabolism, some have found practical applications in medical therapy, wound treatment, and agriculture. These include cycloserine and phosphinothricine, cyclic peptide antibiotics such as gramicidin and bacitracin, chelating peptides (bleomycin), chromopeptides (actinomycin), and depsipeptides (virginiamycin). Valinomycin selectively binds K^+ ions. Some of these peptides are formed on the ribosome, while others are formed by non-ribosomal synthesis. Many peptide antibiotics were isolated from Streptomyces strains, but some were isolated from such Gram-positive microorganisms as Streptococci and Bacilli.

Amino acid derivatives. D-Cycloserine is synthesized by *Streptomyces orchidaceus*. An analog of the bacterial cell wall component D-alanine, it inhibits alanine racemase, a key enzyme in murein biosynthesis. Due to its excellent efficacy against *Mycobacterium tuberculosis*, it was previously the antibiotic of choice, in combination with rifampicin, for treating tuberculosis. Alanyl-alanyl phosphinothricine, first isolated from *Streptomyces hygroscopicus*, is a herbicide ("Bialaphos"). In a plant, it is degraded to phosphinotricine, an analog of L-glutamic acid which inhibits the glutamine synthetase of plants. Phosphinothricine (Glufosinat®, Basta®) (→282) is industrially manufactured by chemical synthesis. If acetyl transferase from *S. hygroscopicus* is cloned and expressed in an agricultural plant, the plant becomes resistant to phosphinothricine while weeds remain susceptible.

Peptide antibiotics are built from 10–50 amino acids with a high proportion of hydrophobic and positively charged side chains, resulting in a strong affinity towards the cytoplasmic membrane. Their biosynthesis occurs either through the ribosome or non-ribosomal enzyme complexes. *Ribosomal peptide antibiotics* begin as linear peptides but later undergo post-translational modifications to non-proteinogenic amino acids, e. g., through the epimerization of L- to D-amino acids or through formation of lanthionine, a cystine analogue that can form peptide chains bridged by thioether groups (lantibiotics). An example is nisin, produced by Lactobacilli, and found in essentially all fermented milk products. It lyses the cytoplasmic membrane of bacteria and thus helps preserve milk products. Defensins (35–45 amino acids) (→118) are another group of cysteine-rich, cationic peptide antibiotics and are part of the immune system: they protect the intestinal mucosa against damage from microorganisms. *Non-ribosomal peptide synthesis* takes place on the thio template of a soluble multi-enzyme complex that resembles the fatty acid synthase of eukaryotes. In this system, short linear peptides are produced, which can be further transformed into cyclic peptides (e. g., Gramicidin S). They rarely contain >20 building blocks and frequently contain unusual amino acids or additional structural elements. Due to toxicity, their use is limited to external applications, such as in the treatment of wounds or burns. Bacitracin is also used as a feed additive. Cyclosporin is synthesized by *Tolypocladium inflatum* and is the immunosuppressant of choice in solid organ and bone marrow transplantation. As it inhibits T-cell activation, it is also used in some chronic inflammatory diseases such as nephrotic syndrome, refractory Crohn's disease, ulcerative colitis, and rheumatoid arthritis. The colistins (polymycines) produced by *Bacillus polymyxa* are important reserve antibiotics against infections by Gram-negative microorganisms. The bleomycins produced by *Streptomyces verticillus* are among the most important anti-tumor antibiotics. Their 1:1 complex with Fe^{3+}, in the presence of triplet oxygen, acts like a DNase and leads to the rupture of single DNA strands. Actinomycin, a peptide derivative of phenoxazinone, is formed by various Streptomyces strains. It intercalates into 5'-TGCA-3' palindrome sequences of DNA, thus blocking translation. The resulting cytotoxicity effects were used for some time in tumor treatments. The amino acid building blocks of the depsipeptides are alternately linked via ester and amide bonds. They are mainly active against Gram-positive bacteria. Virginiamycin, produced by *Streptomyces virginiae*, was used in large quantities in pig and calf fattening. The siderochromes are peptide antibiotics that bind Fe and contain hydroxamic acid groups. They are sometimes used for treatment of iron-storage diseases.

Amino acid derivatives



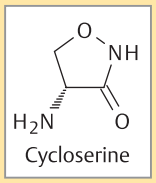
Phosphinothricin

$C_5H_{12}NO_4P$

M_R 181.13

CAS 35597-44-5

(S)-Form

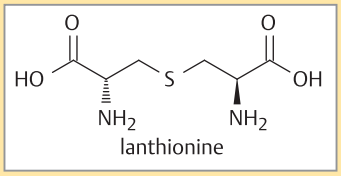


Cycloserine

$C_3H_6N_2O_2$

M_R 102.09

CAS 68-41-7



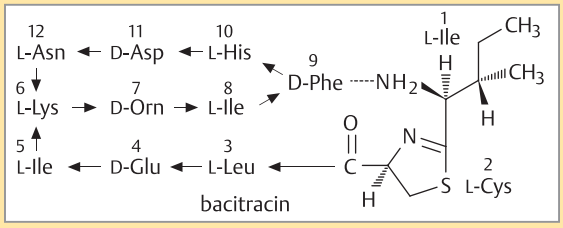
lanthionine

$C_6H_{12}N_2O_4S$

M_R 208.2318

CAS 922-55-4

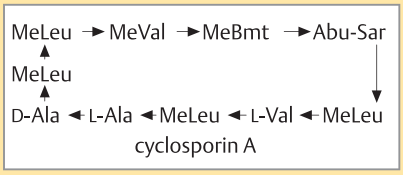
Cyclic peptide antibiotics



bacitracin

$C_{66}H_{103}N_{17}O_{16}S$ M_R 1422.71

very soluble in water CAS 1405-87-4

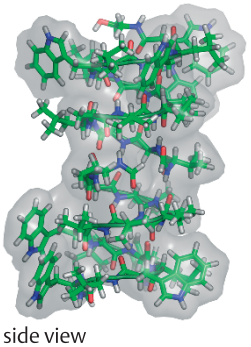


cyclosporin A

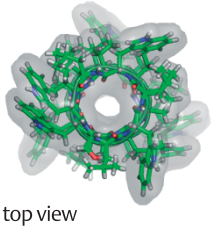
$C_{62}H_{111}N_{11}O_{12}$

M_R 1202.63 CAS 59865-13-3

Abu	L-aminobutyric acid
Sar	sarcosine
MeBmt	butenyl dimethyl threonine
MeLeu	methyl leucine



side view



top view

gramicidin A, A- and B-chain, phospholipid membrane (1MAG)

antibiotic	organism
phosphinothricine	chemical synthesis
bacitracin A	*Bacillus licheniformis*
polymyxin	*Bacillus polymyxa*
gramicidin	*Bacillus* sp.
bleomycin	*Streptomyces verticillus*
cyclosporin	*Tolypocladium inflatum*
virginiamycin	*Streptomyces virginiae*
valinomycin	*Streptomyces fulvisimus*

Fermentation of bacitracin

prefermenter

1 – 3 m^3, 6 h at 37 °C

bioreactor

~ 100 m^3, 30 h at 37 °C; saccharose, soy meal, salts

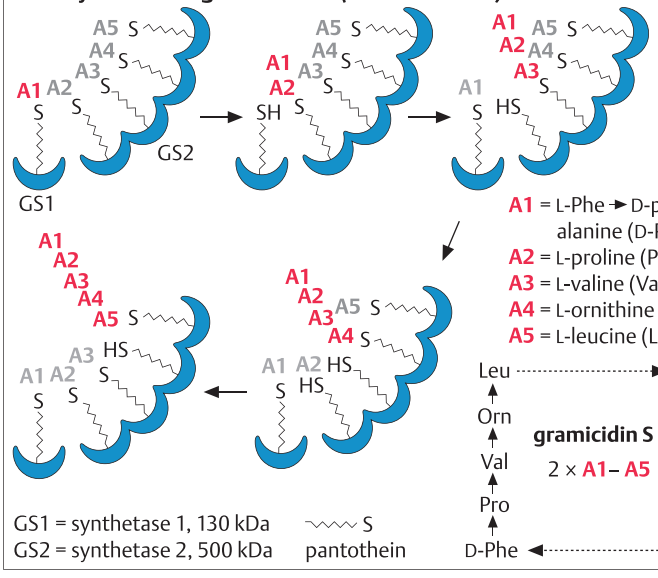
recovery

pharmaceuticals: extraction with 1-butanol; ion exchanger

animal feed: spray drying of fermentation broth

9 g L^{-1} in 30 h

Biosynthesis of gramicidin S (*Bacillus brevis*)



GS1 = synthetase 1, 130 kDa

GS2 = synthetase 2, 500 kDa

S pantothein

Glycopeptide, lipopeptide, polyether, and nucleoside anti-biotics

General. This category includes the very important glycopeptides vancomycin, which is indispensable for the treatment of methicillin-resistant *Staphylococcus aureus* strains (MRSA) (→204) ("first-line treatment"), the acylated glycopeptide teicoplanin and the lipopeptide daptomycin. The polyether antibiotic monensin is used in chicken feed, where it shows prophylactic efficacy against protozoa. Although nucleoside antibiotics occur naturally, only synthetic analogs are used in therapy, e. g., the guanosine analog acyclovir for the therapy of viral meningitis. The glycoside lincomycin is highly effective against Gram-positive enterobacteria.

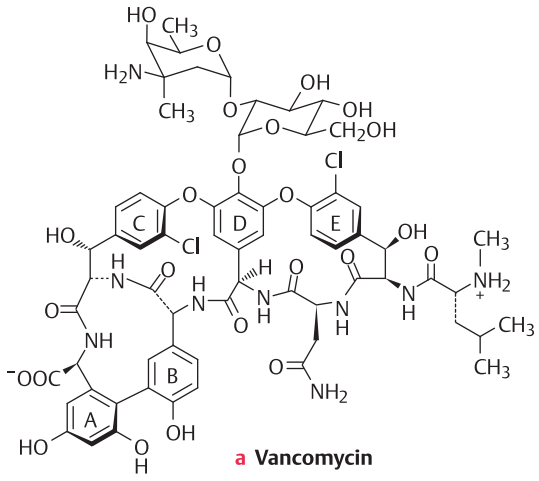
Glyco- and lipopeptide antibiotics. Vancomycin is produced by *Amycolatopsis orientalis*, an Actinomyces strain. It is used against penicillin-resistant Enterococci, e. g., in septic endocarditis, and for patients who are allergic to penicillin. Since it is nephrotoxic and has been or is sometimes used in combination with other nephrotoxic antibiotics, such as aminoglycosides and cyclosporin, thorough monitoring of nephrotoxic side-effects is absolutely required. The effect of vancomycin, like that of β-lactam antibiotics, relies on inhibition of bacterial cell wall biosynthesis (binding to UDP muramyl pentapeptide). Resistant strains form a modified cell wall glycopeptide that does not react with vancomycin. It is assumed that this type of resistance is horizontally transferred via transposons between humans and domestic animals. Teicoplanin is a mixture of glycolipids carrying various acyl side chains. Its broad spectrum of action is similar to that of vancomycin. Daptomycin is a cyclic peptide built from 11 amino acids and a decanoyl side-chain. It is produced by *Streptomyces roseosporus* and is presently (2014) among the safest antibiotics (i. e., to which the least resistance has been observed) and which can be used against infections even when vancomycin has failed. Daptomycin forms pores in bacterial cell membranes, resulting in the efflux of K^+ ions and a depolarization of the membrane. Like vancomycin it must be applied parenterally. Cyclosporin A or cyclosporin is another cyclic lipopeptide. It is synthesized by fungi of the genus *Beauveria*. It acts as an immunosuppressant by a complex mechanism preventing the activation of lymphocytes, and has revolutionized organ transplantations. It is also used for the treatment of skin and intestinal diseases.

Lincomycin, an antibiotic formed by *Streptomyces lincolnensis*, is active against Gram-positive pathogens and is used in veterinary medicine. Similar to chloramphenicol, it binds to the 50S subunit of the ribosome, inhibiting extension of the growing peptide chain. Resistant strains occur frequently. They either produce rRNA modified by methylation or detoxify lincomycin by enzymatic transformation.

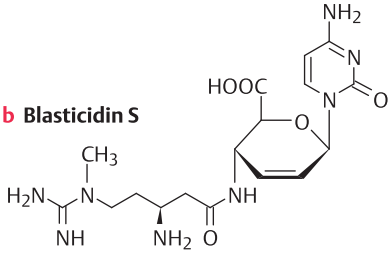
Monensin, a polyether antibiotic, is manufactured by fermentation of *Streptomyces cinnamoensis*. Biosynthesis proceeds via polyketides, using acetate, propionate, and butyrate as building blocks. Monensin incorporates as a ionophore into membranes and causes osmolysis of cells through the influx of Na^+ ions. This mechanism of action affects not only bacteria and fungi, but also protozoa, e. g., *Eimeria* sp. and *Toxoplasma* sp., which occur during mass production of domestic animals. Although its toxicity is high in humans and horses (but not in chicken and cattle), it has become one of the most important broad-spectrum antibiotics in chicken feeds and is well tolerated if dosed appropriately. It is registered both in the European Union and in the USA. Its market share for this application is 80 %, counting also salinomycin, which is structurally related. It is produced in quantities of several thousand tons. Care must be taken not to poison other farm animals such as horses, or farm workers when using it.

Nucleoside antibiotics have found only limited applications. The cytosine analog blasticidin S from *Streptomyces griseochromogenes* was used as a fungistatic agent for the treatment of blight in rice agriculture. It prevents the binding of aminoacyl tRNA to the ribosome. It is quite toxic to plants as well as to fish, animals, and man, affecting the mucous membrane, skin, and lungs, and was replaced by Kasugamycin. A purine derivative important in medical therapy is acyclovir, a synthetic guanosine derivative. It is active against Herpes virus and can be used against viral encephalitis. It is obtained by chemical synthesis.

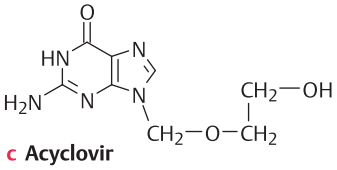
Glycopeptide, glycoside, and nucleoside antibiotics



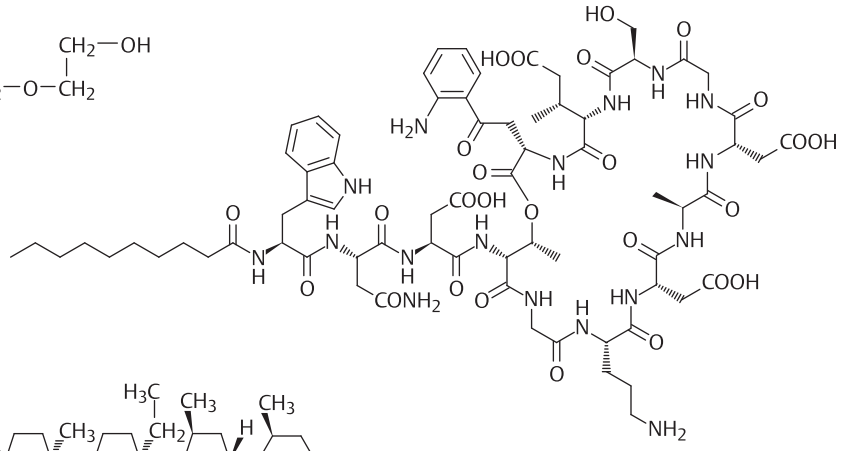
a Vancomycin



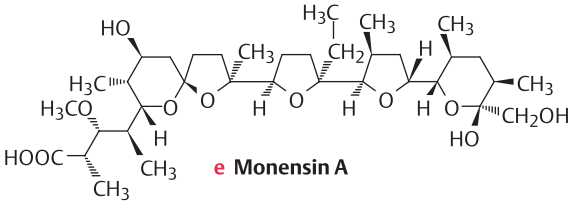
b Blasticidin S



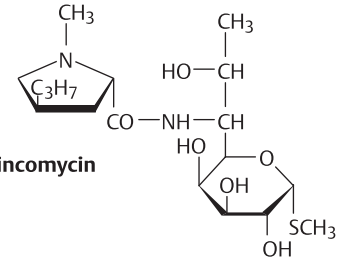
c Acyclovir



d Daptomycin



e Monensin A



f Lincomycin

	a	b	c	d	e	f
	$C_{66}H_{75}Cl_2N_9O_{24}$	$C_{17}H_{26}N_8O_5$	$C_8H_{11}N_5O_3$	$C_{72}H_{101}N_{17}O_{26}$	$C_{36}H_{62}O_{11}$	$C_{18}H_{34}N_2O_6S$
M_R	1449.27	422.44	225.21	1619.7086	670.90	406.56
CAS	1404-90-6	2079-00-7	59277-89-3	103060-53-3	17090-79-8	154-21-2

antibiotic	produced by	type	applications
a Vancomycin	*Amycolatopsis orientalis*	glycopeptide	MRSA
b Blasticidin S	*S. griseochromogenes*	nucleoside	plant protection fungicide
c Acyclovir	chemical synthesis	nucleoside	clinical antiviral agent
d Daptomycin	*S. roseosporus*	lipopeptide	MRSA
e Monensin A	*S. cinnamonensis*	polyether	coccidiosis prevention in animal breeding
f Lincomycin	*S. lincolnensis*	lincosamide	veterinary medicine

S. = *Streptomyces* MRSA = methicillin-resistant *Staphylococcus aureus* infections

Aminoglycoside antibiotics

General. The discovery of streptomycin by Selman Waksman (1943) was a milestone in the development of antibiotics. Streptomycin permitted, for the first time in history, a treatment of tuberculosis, caused by *Mycobacterium tuberculosis*. Today, the nephrotoxic properties of streptomycin (which are typical of aminoglycoside antibiotics) have led to its replacement by isonicotinic acid hydrazide and rifampicin (formerly by cycloserine). Due to the unusual structure of the mycobacterial cell wall, which hinders the entry of drug, treatment is difficult, and active TB requires the prolonged use of antibiotic combinations. The lead structure of most aminoglycoside antibiotics is an aminocyclitol ring, i. e., 2-deoxy streptamine, which is linked by glycosidic bonds to other amino sugars (→32). Aminoglycoside antibiotics exhibit broad biological efficacy and are also active against many Gram-negative pathogens. As a result, they are the antibiotics of choice for severe infections and thus hold a firm place in human therapy, in spite of their high toxicity (they must be parenterally applied) and the formation of resistant strains. They are also used in plant protection. Their sales in the world market are ca. 1 billion US$ (2009). The most important compounds for human therapy are the gentamicins, the neomycins, tobramycin, kanamycin, and semi-synthetic products such as sisomicin and amikacin. Streptomycin is occasionally still used. Spectinomycin has been successfully used against penicillin-resistant *Neisseria gonorrhoeae* infections. Kasugamycin is an important agricultural antibiotic for combating rice blight. Hygromycin was used as an antihelminthic in pig and poultry fattening, but this use is no longer permitted in the EU as cross-resistances to aminoglycoside antibiotics used in human therapy have been observed.

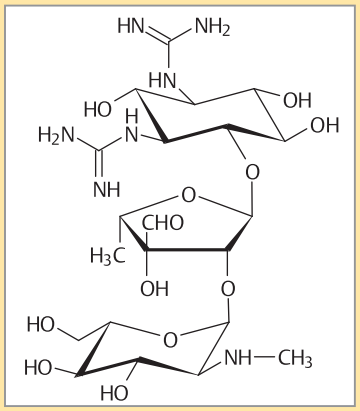
Biosynthesis. The genera Streptomyces and Micromonospora mainly form the aminoglycoside antibiotics. The multi-step biosynthesis (streptomycin: 24 steps) always starts from D-glucose and usually leads, on the scaffold of nucleotide sugars, to an amincyclitol unit which is glycosidically linked to other unusual sugars (amino sugars, C-branched sugars). L-streptose is a branched-chain hexose of unusual composition (5-deoxy-3-C-formyl-L-lyxose); it is formed from D-glucose via L-rhamnose through deoxythymidinediphosphate-linked activated intermediates. The biosynthesis of streptomycin requires 33 proteins, and all genes required for this pathway are located on a single gene cluster of the *Streptomyces griseus* genome. They comprise 30–40 bp and have mostly been cloned. However, metabolic regulation of the biosynthetic pathway is not yet completely understood.

Production. As with other industrially manufactured antibiotics, the production strains have undergone significant strain improvement by series of mutation, selection, and back-crossing. Using such procedures, the antibiotic yield of the wild-type strain, which is on the order of some mg L^{-1}, has been raised more than 1000 fold (streptomycin: > 10 g L^{-1} after 120 h fermentation). Industrial production is carried out in large bioreactors, using glucose, starch, or dextrin as the carbon source and soymeal as the nitrogen source. Aminoglycosides that tend to bind to the mycelium, such as gentamicin, are released by acidification to pH 2.0. After separation of the cell mass by filtration or centrifugation and concentration of the fermentation broth, the antibiotic is purified by several cycles of ion-exchange chromatography and crystallization.

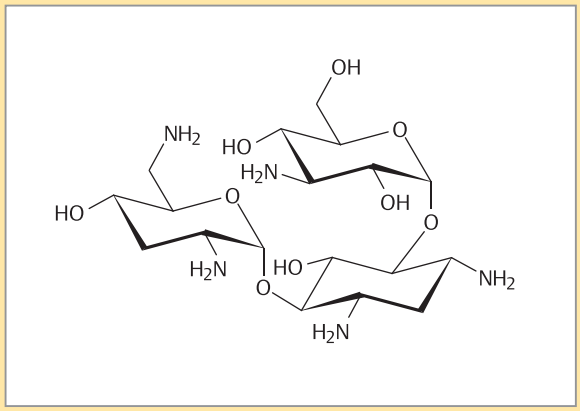
Mechanism of action and resistance. Aminoglycoside antibiotics bind to the 30S subunit of the ribosome and interfere with translation, eventually resulting in the inhibition of protein biosynthesis (→202). Some of them also bind specifically to type-I introns of RNA. A major disadvantage is that they lead quickly to the formation of resistant phenotypes, which can be coded on plasmids or on the chromosome. Resistant strains have either mutated the aminoglycoside binding site on their ribosomes, or have acquired the capacity, usually by transfer of plasmids carrying the respective genes, to block key hydroxyl groups via acetylation, phosphorylation, or adenylation.

Semisynthetic aminoglycosides of high potency can be obtained by chemical derivatization, preferentially at amino groups (→220). Examples are sisomicin, amikacin, and tobramycin. Experiments to achieve semisynthetic aminoglycosides through pathway engineering (combinatorial biosynthesis), however, have led to only minor success so far, although in many cases all genes of a pathway have been cloned and suitable expression cassettes have been prepared. It is assumed that many of the enzymes in such a pathway are subject to complex individual regulation, rendering pathway interchanges difficult.

Aminoglycoside antibiotics



streptomycin M_R 581.58
$C_{21}H_{39}N_7O_{12}$ CAS 57-92-1

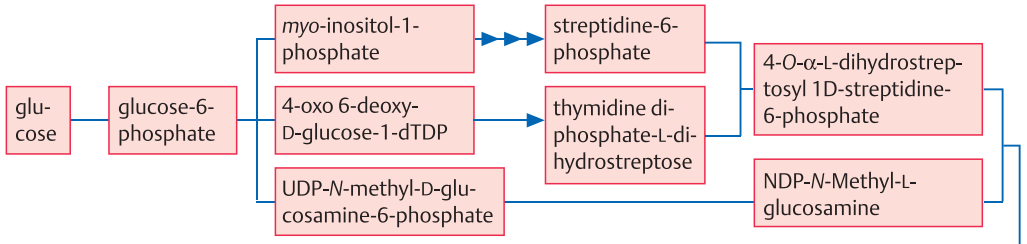


tobramycin M_R 467.51
$C_{18}H_{37}N_5O_9$ CAS 32986-56-4

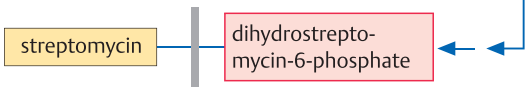
Antibiotic	Organism	Applications
Streptomycin	*Streptomyces griseus*	endocarditis, tuberculosis
Kanamycin	*S. kanamyceticus*	local infections (skin, eye)
Tobramycin	*S. tenebrarius*	infection of cystic fibrosis patients
Gentamicine	*Micromonospora purpurea*	infections with Gram-negative pathogens
Netilmicin*	*M. purpurea*	infections with Gram-negative pathogens
Amikacin*	*Streptomyces kanamyceticus*	infections with multi-resistent pathogens
Neomycine	*S. fradiae*	local infections (skin)
Kasugamycin	*S. kasugaensis*	blight infections of rice
Validamycin	*S. hygroscopicus*	rice agriculture

*chemical derivative

Biosynthesis of streptomycin



for this biosynthetic pathway, a total of 33 enzymes and other proteins is required. The genes are arranged in a gene cluster of 30 – 40 kbp and have been largely cloned



cytoplasmic membrane

Manufacture

preculture	fermentation	removal of cells	purification	
inoculation and growth	up to 150 m^3 bioreactor, fed-batch mode, 28 – 30 °C , 70 – 120 h, glucose or dextrins, soy meal, 1 – 3 g L^{-1} NaCl, trace elements, 0.5 – 1 vvm O_2	for mycelium-bound antibiotic, acidify to pH 2	ion exchangers, crystallization and recrystalli-zation	> 10 g L^{-1} after 120 h

Tetracyclines, chinones, chinolones, and other aromatic antibiotics

General. Due to their activity against a broad range of pathogens, the tetracyclines are an important class of antibiotics. They are used in medicine and agriculture, and their market volume in 2009 was ~1.6 billion US$. Fluorochinolones are synthetic analogs of nalidixic acid. They are prepared by chemical synthesis and constitute the major class of medical antibiotics after the lactams (market value 2009: 7.1 billion US$).

Tetracyclines. First described in 1945 as metabolites of *Streptomyces aureofaciens,* tetracyclines have become very important antibiotics due to their low toxicity and broad activity spectrum. They are active against Gram-positive and Gram-negative bacteria, Rickettsia, Mycoplasma, Leptospira, Spirochaeta, and some larger viruses. Unfortunately, tetracyclines are widely used in some countries as feed additives for chicken and pig fattening, which has led to the development of resistant strains (→204). Mechanisms of resistance that have been observed often are: reduced penetration of the antibiotic through the outer membrane of Gram-negative cells (altered porins), modification of their ribosomal binding site, and plasmid-coded synthesis of so-called tet-proteins that support rapid export of tetracyclines from the bacterial cell (→204). These resistance mechanisms have more seldom been observed with the use of semisynthetic tetracyclines such as doxycycline or tigecycline. Tetracyclines inhibit protein biosynthesis by binding to the 70S ribosome (→202). They are formed exclusively by Streptomyces strains; the primary product is usually oxytetracycline. Biosynthesis from glucose requires many steps and passes through polyketide intermediates. Thus, the gene cluster for oxytetracycline synthesis in *S. rimosus* comprises at least 21 genes, including three genes for antibiotic resistance. For industrial manufacturing, optimized production strains of several Streptomyces species are cultivated in large bioreactors. Yields may reach >80 g L^{-1} if the O_2 supply is optimized and the phosphate content of the medium is well controlled. For isolation, the cell mass is separated, and the broth is extracted using n-butyl acetate. Ion-exchange chromatography or precipitation with chaotropic reagents are used for down-stream processing.

Anthracyclines. Anthracycline glycosides such as doxorubicin (adriamycin) inhibit DNA replication through intercalation and inhibition of topoisomerases. They are used in chemotherapy for tumor treatments and are produced by fermentation.

Fluorochinolones. Nalidixic acid is a side-product of the chemical synthesis of chloroquinone, an anti-malaria agent. Its bactericidal effect was discovered in 1962 and shown in 1977 to be due to the inhibition of bacterial topoisomerase (gyrase A subunit). Since the structure and function of bacterial and human topoisomerases differ quite significantly, chinolones exhibit low toxicity to humans. On the other hand, they show a wide activity spectrum against Gram-positive and -negative bacteria, Mycobacteria, Chlamydia, and anaerobic microorganisms. Microbial resistance develops only slowly and is not plasmid-coded. If developed, it is due to modifications of the gyrase subunit or to reduced membrane permeation. Among the >5000 chinolone derivatives, which are exclusively produced by chemical synthesis, some have found wide application. Ciprofloxacin (Ciprobay®) is a chinolone antibiotic active against *Bacillus anthracis*.

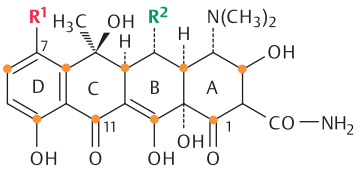
Chloramphenicol was isolated in 1950 from cultures of *Streptomyces venezuelae* but today is completely produced by chemical synthesis. Its biosynthetic pathway branches from the biosynthesis of aromatic amino acids at the level of chorismic acid. It acts through binding to the 50S subunit of 70S ribosomes, blocking peptidyl transferase. Chloramphenicol was the first broad-spectrum antibiotic. It is effective against a wide range of Gram-positive and -negative bacteria, Actinomycetes, Rickettsia, and large viruses. Since it damages the bone marrow, it is only used as a reserve antibiotic, in particular for the treatment of typhus, shigellosis, and Rickettsia-based infections.

Griseofulvin, a benzophenone derivative, is a fungistatic antibiotic that blocks mitosis through binding to the spindle apparatus of fungi, which is detected by the formation of short, curly hyphae. It is produced by fermentation. Applications include the treatment of dermatomycosis and agricultural use as a leaf fungicide against blights.

Chinoid and aromatic antibiotics

antibiotic	production	application
(chloro-)tetracycline oxytetracycline	*Streptomyces aureofaciens* *Streptomyces rimosus*	broad-spectrum antibiotics for human therapy and animal feed
anthracycline	*S. peucetius*	cancer therapy
chinolones	chemical synthesis	broad-spectrum antibiotics for human therapy and animal feed
chloramphenicol	chemical synthesis	broad-spectrum antibiotic
griseofulvin	*Penicillium griseofulvium*	fungal infections on skin or leaves (blight)

tetracyclines

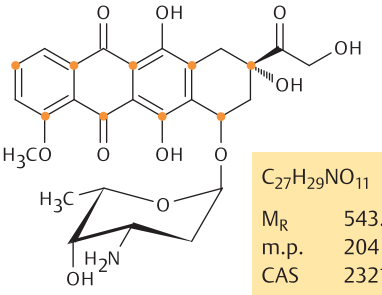


• C^1 of acetate

chlorotetracycline
$C_{22}H_{23}ClN_2O_8$
M_R 478.89
CAS 57-62-5

	R^1	R^2
tetracyline	H	H
chlorotetracycline	Cl	H
oxytetracycline	H	OH

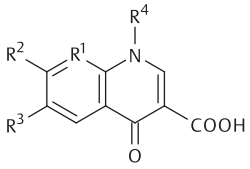
doxorubicin



• C^1 of acetate

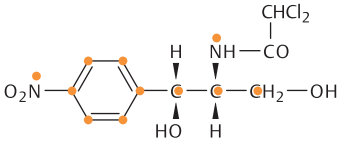
$C_{27}H_{29}NO_{11}$
M_R 543.53
m.p. 204 – 205 °C
CAS 23214-92-8

nalidixic acid and ciprofloxacin, two chinolone antibiotics



	R^1	R^2	R^3	R^4
nalidixic acid	N	CH_3	H	CH_3
ciprofloxacin	CH	$C_4H_9N_2$	F	C_3H_5

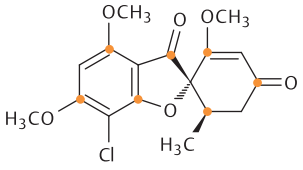
chloramphenicol



• from p-aminophenylalanine

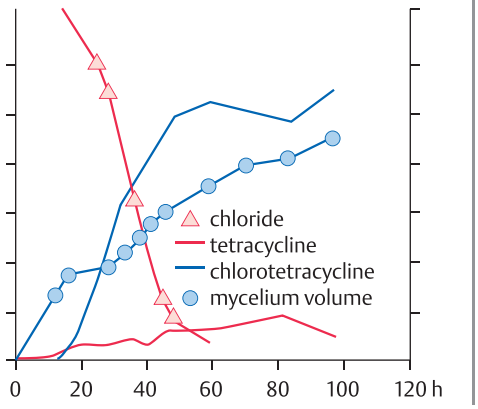
$C_{11}H_{12}Cl_2N_2O_5$ CAS 56-75-7
M_R 323.13

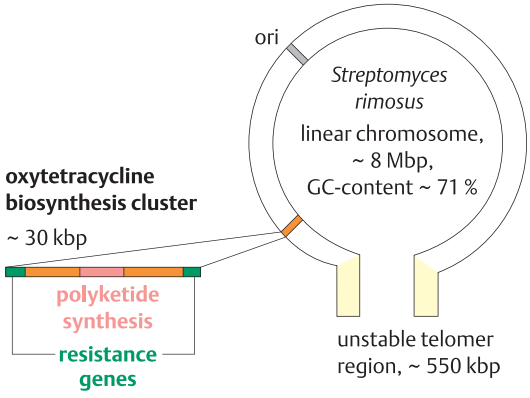
griseofulvin



• C^1 of acetate

$C_{17}H_{17}ClO_6$ CAS 126-07-8
M_R 352.77





Fermentation and isolation of chlorotetracycline

preculture	bioreactor	isolation	
high-performance strains *Streptomyces aureofaciens*	~ 150 m^3, saccharose, corn steep liquor, salts, 1 vvm air; 28 °C, pH 5.8 – 6.0; 60 – 65 h	separation of mycelium through filter press or separator, multi-stage extraction with n-butyl acetate, purification by ion-exchange chromatography/salt precipitation	>80 g/L after 140 h

Macrolide antibiotics

General. This important group of antibiotics includes the macrolide, the polyene, the macrotetrolide, and the ansamycine antibiotics. They all share a macrocyclic lactone or lactam ring that is formed from a long-chain polyhydroxy fatty acid and a terminal hydroxyl or amino group. The ring can be glycosylated with unusual sugars (macrolides), it may contain conjugated double bonds (polyenes) or an aromatic chromophore (ansamycines), or it may be built as a polylactone (macrotetrolides). Most polyketides are isolated from Streptomyces strains. They are used in human therapy, and also for food protection and in animal feeds. The market value of macrolides used in medicine is in the range of 5 billion US$ (2009) or about ⅕ of the total antibiotic production.

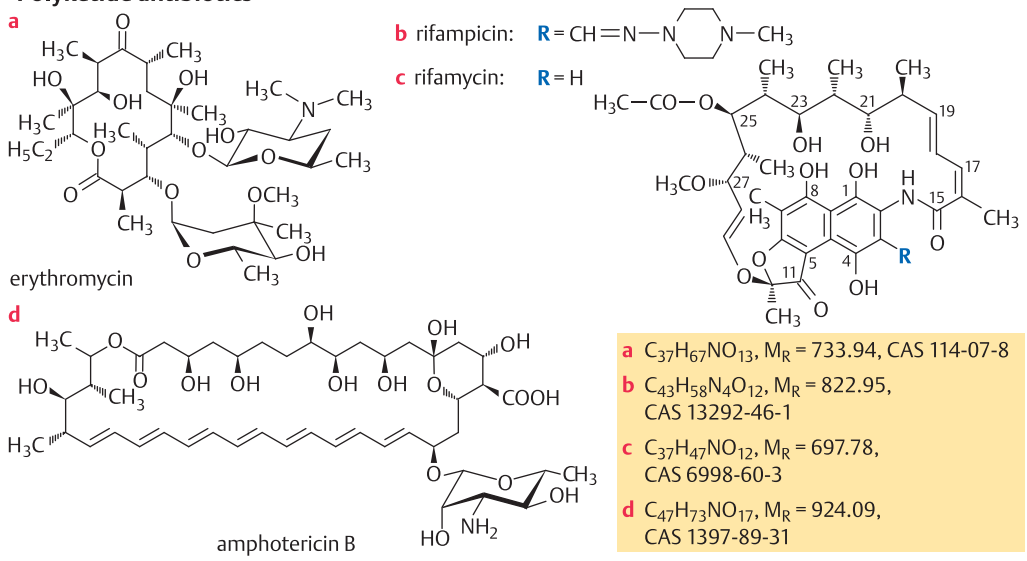
Macrolide antibiotics are lipophilic, often basic, compounds. The key structural element is a 10–60-membered macrocyclic lactone ring, which is formed by a multi-enzyme complex resembling fatty acid synthase, through condensation of an acyl-CoA starter unit with malonyl-CoA or its methyl and ethyl homologs. Polyketides are the hypothetical intermediates of this reaction, and they are further modified by unusual sugars containing, for example, amino groups, C-methyl branches, and deoxy groups. Macrolide antibiotics exhibit low toxicity and thus are often used in pediatric medicine. They preferentially inhibit Gram-positive microorganisms, binding to their 50S ribosomal subunits and interrupting translocation of the growing peptide chain. Formation of resistant pathogens is frequently observed and mostly due to methylation of the ribosomal 23S RNA. Azithromycin, Clarithromycin, Erythromycin and Spiramycin are preferred antibiotics against bacterial infections of the respiratory tract. A new representative of this group is Cethromycin. Tylosine, a related macrolide, was a valuable feed antibiotic for pig fattening, due to its activity against mycoplasms. Since this application led to the development of cross-resistant strains (→204), the use of tylosine in the European Union requires prescription by a veterinarian.

Polyene antibiotics are formed preferentially by Streptomyces strains. They are constituted from 26- to 38-membered macrocyclic lactone rings with 3–7 conjugated double bonds and may contain additional building blocks such as amino sugars in glycosidic linkage. Several polyene antibiotics are used as fungistatic agents, e. g., amphotericin B or nystatin for medical therapy of *Candida albicans* and pimaricin (natamycin) as a food preservative in cheeses. Polyene antibiotics function by complexing with fungal membrane sterols such as ergosterol. As a result, they are not effective against bacteria. Since they are nephro- and hepatotoxic, their use is limited to severe infections. They are too labile to be used as fungicides in agriculture.

Ansamycins are macrolactam antibiotics containing an aromatic chromophore. Their most important representative is rifamycin, a product of *Amycolatopsis mediterranei*. It is highly active against Gram-positive bacteria and Mycobacteria. Its biosynthesis involves the assembly of a polyketide by addition of acetate and propionate chain extension units to a unique starter unit, 3-amino-5-hydroxy-benzoic acid (AHBA), which is formed by a pathway similar to shikimic acid formation. Rifampicin, a semisynthetic derivative of rifamycin, is presently the most important agent for the treatment of tuberculosis (pathogen: *Mycobacterium tuberculosis*) and is also applied in the treatment of leprosy (*Mycobacterium leprae*) or *Legionella pneumophila* infections. It inhibits the bacterial DNA-dependant RNA polymerase by binding to its β-subunit. Since rifampicin does not bind to human RNA polymerase, it is nontoxic to humans. In resistant strains, the RNA polymerase is modified through mutations.

Fermentation and purification. The commercially important macrolide antibiotics are produced industrially by high-performance strains in large bioreactors. Complex media are used, and propionic acid is added as an important precursor for biosynthesis. Yields of erythromycin by *Saccharopolyspora erythraea* (former name: *Streptomyces erythreus*) are in the range of 10 g L^{-1} after 72 h. After removing the cell mass by filtration or separators, the extracellular products are isolated via solvent extraction (→104), for instance with butyl acetate, and further purified by chromatography and recrystallization.

Polyketide antibiotics



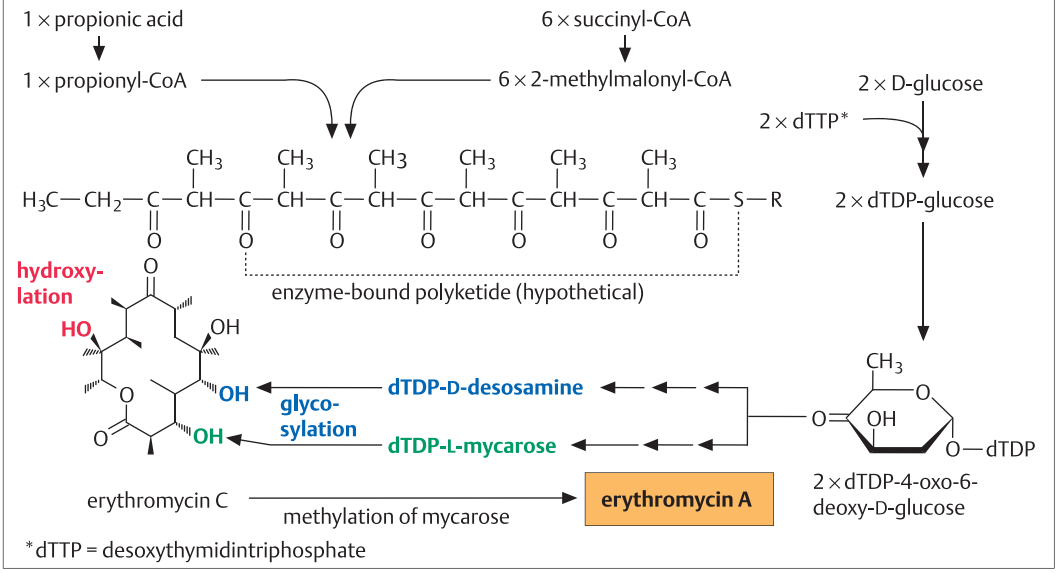
a $C_{37}H_{67}NO_{13}$, M_R = 733.94, CAS 114-07-8

b $C_{43}H_{58}N_4O_{12}$, M_R = 822.95, CAS 13292-46-1

c $C_{37}H_{47}NO_{12}$, M_R = 697.78, CAS 6998-60-3

d $C_{47}H_{73}NO_{17}$, M_R = 924.09, CAS 1397-89-31

Biosynthesis of erythromycin A



*dTTP = desoxythymidintriphosphate

Manufacture

erythromycin	rifamycin/rifampicin
fermenter	**fermenter**
high-performance strains of *Saccharopolyspora erythraea*, bioreactor > 120 m^3 in fed-batch mode, glucose, soy meal, trace elements, 0.2 – 0.5 % propionic acid; 33 °C, 70 – 120 h	high-performance strains of *Nocardia mediterranei*, bioreactor > 120 m^3 fed-batch mode; glucose, soy meal, trace elements, 0.2 – 0.5 % propanol; 33 °C, 70 – 120 h

removal of cells by filtration or separators

purification

counter current extraction with acetic acid butyrate, chromatography, recrystallization

~ 10 g L^{-1} erythromycin after 72 h	~ 7 g L^{-1} rifamycin after 72 h
	4 step Mannich synthesis to rifampicin

New pathways to antibiotics

General. Although antibiotic therapy has been a success story, and new antibiotics are isolated year after year, infections that were believed extinct do reappear and prove more difficult to treat. The reemergence of tuberculosis is just one of these problems, and even treating infections from Staphylococci or Streptococci today may be difficult. At the root of this problem lies bacterial antibiotic resistance against one or several antibiotics ("multi drug resistance"), which constitutes a major medical and scientific challenge (→204). Antibiotic resistance is often coded on plasmids or transposons and can be horizontally transferred among different microorganisms, e. g., by conjugation or phage infections. Another critical problem is the relatively small number of antibiotics that are effective against pathogenic fungi and yeasts that are nontoxic to man. Since the metabolisms of eukaryotic organisms are closely related, antibiotics that affect fungi are usually toxic against man as well. As a consequence, there are good reasons to look for new leads in the discovery of antibiotics.

New screening procedures. If standard screening procedures, e. g., bioassays, are used for the discovery of new antibiotics, nine out of ten hits turn out to provide structures that have already been described. As a result, a range of unconventional procedures has been developed in order to discover new lead structures. The most important examples are: 1) precursor-directed biosynthesis, where the addition of synthetic precursors to the fermentation leads to biosynthetic antibiotics, 2) screening for antibiotics in hitherto neglected genera such as myxobacteria, rare actinomycetales, lichens, sponges, or microorganisms which live in symbiosis with plants or animals (endophytic microorganisms), 3) modifications of the screening procedure, relying on novel assays, 4) genetic recombination of strains which produce antibiotics, and 5) combinatorial biosynthesis by *in vitro* recombination of genes participating in antibiotic synthesis.

Reverse genetics. Due to the great advances in sequencing genes and genomes, many gene clusters coding for antibiotic products are already known. If enough gene sequences are known for those enzymes that carry out key steps in antibiotic synthesis, e. g. in hydroxylation, glycosylation or halogenation, their genetic footprint (consensus sequences) can be used to screen for similar enzyme activities in the genomic DNA of other organisms or in metagenomes. As an example, this method was successfully used for the screening of new halogenases.

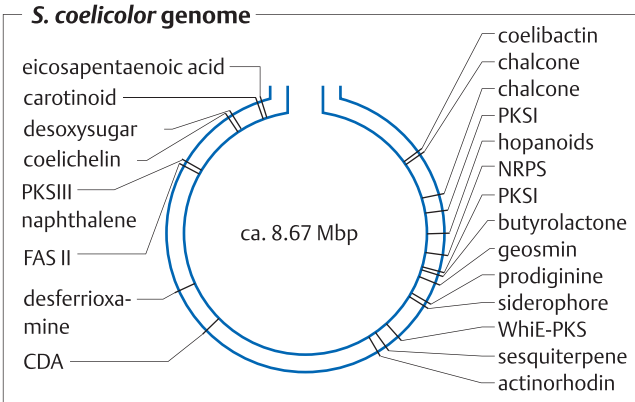
Combinatorial biosynthesis. The biosynthesis of most antibiotics is coded in gene clusters (modules) which comprise a) structural genes for the required enzymes, b) genes imparting resistance to the producing organism, c) genes that regulate product formation, and d) genes important for the export of the antibiotic. Many of such modules are documented in antiSMASH (**anti**biotics and **S**econdary **M**etabolites **A**nalysis **Sh**ell), a public database, and can be accessed and used for random or "semi-random" gene shuffling experiments (→198) (in the latter case, a preselection of gene function is done before shuffling). Streptomyces strains are an ideal host for such experiments, since they each contain 20–30 putative gene clusters for the biosynthesis of antibiotics, and during conjugation among two strains a relatively high plasmid transfer takes place, which can be exploited for combinatorial biosynthesis using such modular "gene bricks" (*"in vivo* synthetic biology") (→320). Using this procedure, new antibiotics that do not occur in nature have already been synthesized.

New targets from genome analysis. The number of pathogenic microorganisms whose genome has been fully sequenced is growing year by year. Thus, the genomes of *Haemophilus influenzae* (1.83 Mbp; chronic bronchitis), *Helicobacter pylori* (1.67 Mbp; ulcers), *Borrelia burgdorferi* (0.91 Mbp; Lyme disease), *Mycobacterium tuberculosis* (4.41 Mbp; tuberculosis), *Treponema pallidum* (1.14 Mbp; syphilis) and *Chlamydomonas trachomatis* (1.04 Mbp: eye infections) have all been sequenced and analyzed. It was assumed that metabolic or signal-transduction pathways specific for the pathogen can be deciphered from genomic information, leading to pathogen-specific targets which can be used in drug development. Unfortunately, in a seven-year industrial study using 300 putative targets and millions of drug candidates in a high-throughput screening format, only five suitable lead structures were identified.

Precursor-directed biosynthesis and mutasynthesis (examples)

type of antibiotic	producer	precursor fed to the medium
penicillins	*Penicillium chrysogenum*	many aliphatic, alicyclic and aromatic carboxylic acids
cephalosporins	*Acremonium chrysogenum*	*S*-carboxymethyl-L-cysteine
bleomycins	*Streptomyces* sp.	modified amines
bacitracins	*Bacillus licheniformis*	D-allo amino acids
streptomycin	*Streptomyces griseus*	2-deoxy streptidine

S. coelicolor genome



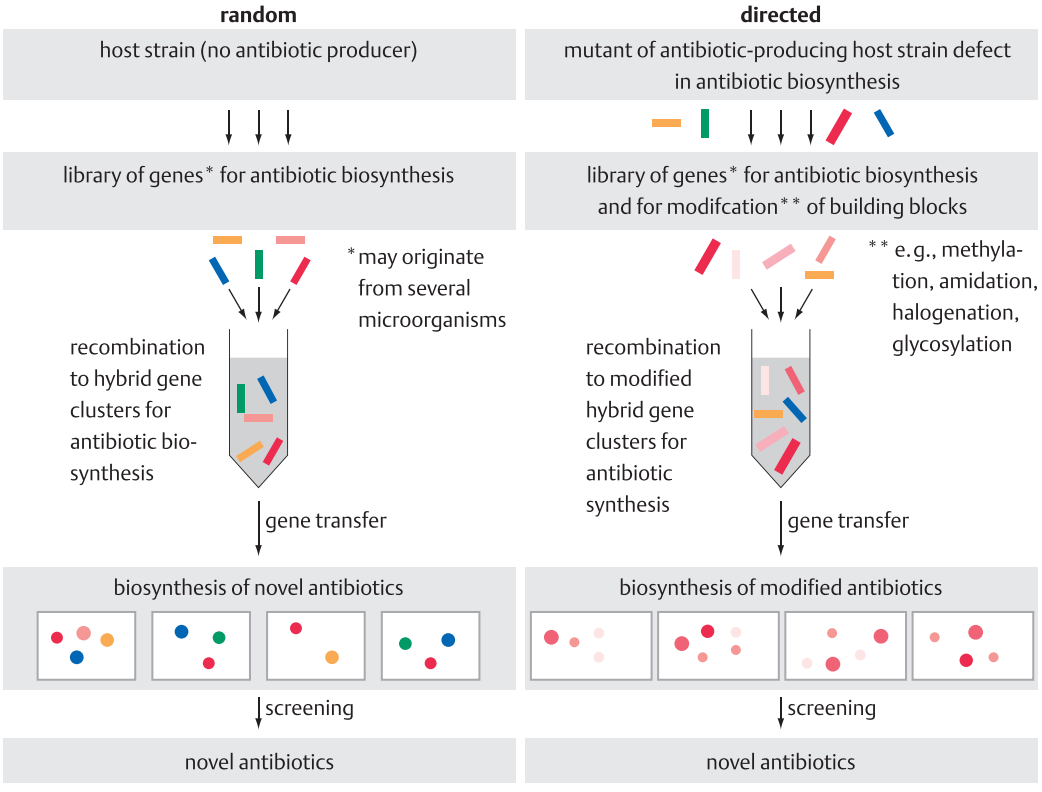
contains > 20 gene cluster for the formation of natural compounds, e.g., antibiotics such as actinorhodine. Additional antibiotics such as methylenomycin are plasmid-coded

Biosynthetic potential of some Actinomycetes strains

Streptomyces coelicolor	8.7	22
Streptomyces avermitilis	9.0	30
Sorangium cellulosum	13.0	17
Saccharopolyspora erythrea	8.2	25
Streptomyces griseus	8.5	34
Salinispora tropica	5.2	17

genome size (Mb)
putative antibiotic-coding gene clusters

Combinatorial biosynthesis



Insulin

General. Insulin is a polypeptide hormone (→28) that regulates the glucose level in the blood of vertebrates. It is also a key drug for treating diabetes mellitus (hyperglycemia). Until 1982, insulin was prepared by extraction from pancreatic glands of slaughtered animals. Since then, recombinant human insulin produced in *Escherichia coli* and *Saccharomyces cerevisiae* has become the dominant production technology. The global market volume was 16 billion US$ in 2010, and the metric production volume is currently an estimated 100 t/y.

Diabetes mellitus is characterized by defects in the synthesis and release of insulin. In type I diabetes, no insulin is formed, due to a genetic defect, virus infection, or autoimmune disease. As a result, the glucose level in the blood must be controlled by a steady supply of insulin via transcutaneous or intramuscular injection. The more common case of type II diabetes (adult-onset diabetes) is mostly a "lifestyle disease". Insufficient production of insulin by the pancreas can sometimes be stimulated by medications, but at a later stage insulin is applied. About 400 million people (2012, world) suffer from diabetes (in the USA: 26 million (2012)), of which < 10 % are type I patients.

Biosynthesis. Insulin is synthesized in the β-cells of the pancreas as preproinsulin, then it is processed to proinsulin, and stored in the Golgi system. When triggered by a complex mechanism involving hypoglycemia, it is hydrolyzed by membrane-bound proteases into 3 polypeptide chains (A, B, and C). The A and B chains (21 and 30 amino acids, respectively) combine via 3 cystine bridges to form active insulin, and the C chain (31 amino) acids is liberated and catabolized.

Production. Insulin therapy was introduced in 1928. For *conventional production*, cattle or pig pancreas was extracted with 1-butanol. The insulin was precipitated as its Zn salt, which is easy to crystallize, and was further purified by gel chromatography ("single-peak insulin"). The amount of insulin that can be prepared from the pancreas of a pig covers the needs of a diabetic patient for 3 days, from a cow for 10 days, resulting in bottlenecks in its industrial production. In addition, human, porcine, and bovine insulin differ in one or two amino acids. As a result, continuous medication of diabetes patients with animal insulin has occasionally resulted in adverse allergic reactions. Although the *chemical synthesis* of human insulin succeeded as early as 1964 in China and Germany, it proved economically unviable. A temporary solution to the allergies arising from the continuous use of animal insulin was found in 1975 by Novo-Nordisk Industri in Denmark, when porcine insulin was transformed into human insulin by *enzyme catalysis* using carboxypeptidase Y, exchanging the C-terminal ala^{30} with thr^{30}. Since 1982, however, the production of *recombinant human insulin by fermentation* has become the method of choice. The DNA encoding preproinsulin was prepared by chemical synthesis (→54), allowing for optimization of the codon usage (→40) with respect to the host organism *E. coli* K12. In an early procedure, the A and B chains were, for safety reasons, expressed and purified separately, and then chemically transformed to the active insulin molecule in an oxidative step. Today, recombinant proinsulin is produced as a fusion protein with tryptophan synthase, and processed by several steps into active insulin. Optimized production strains of *E. coli* synthesize up to 40 % of their cell mass as proinsulin fusion protein. Thus, a 40 m^3 bioreactor provides ca. 100 g of pure recombinant human insulin. A different process starts from the expression of a shortened "mini-proinsulin" using recombinant strains of bakers' yeast. A diabetic can apply insulin using a single-dose "insulin pen". For diabetes type I, portable insulin pumps with catheters are mostly used.

Modern types of insulin. Human insulin with an inverted sequence of $lys^{28}pro^{29}$ (Humalog®, Liprolog®) obtained by protein engineering, is more rapidly available upon injection and thus facilitates the planning of food intake. Faster action is also observed with Insulin Aspart. Insulin glargin (Lantus®) with $pro^{28}asp$ – $thr^{30}arg^{31}arg$ (B chain) and $asn^{21}gly$ (A chain) exhibits prolonged activity, allowing for a less frequent dosage ("single dose insulin"), as does Insulin detemir ($Thr^{30}{}_{del}Lys^{29}$) acylated with a fatty acid at lys^{39}. In insulin degludec, thr^{30} (B-chain) is deleted, and hexadecanedioic acid is conjugated via a γ-L-glutamyl spacer to lys^{29}, resulting in the formation of a multi-hexamer insulin depot in subcutaneous tissue and a slow-release action.

Primary structure of human insulin

A chain

gly - ile - val - glu - gln - cys - cys - thr - ser - ile - cys - ser - leu - tyr - gln - leu - glu - asn - tyr - cys - asn
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21

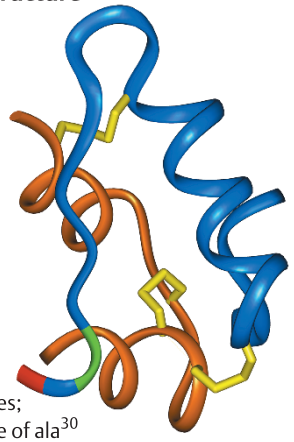
B chain

phe - val - asn - gln - his - leu - cys - gly - ser - his - leu - val - glu - ala - leu - tyr - leu - val - cys - gly - glu
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21

thr ← lys ← pro ← thr ← tyr ← phe ← phe ← gly ← arg
30 29 28 27 26 25 24 23 22

C-terminal thr can be cleaved with **carboxypeptidase Y**, the octapeptide with **trypsin**

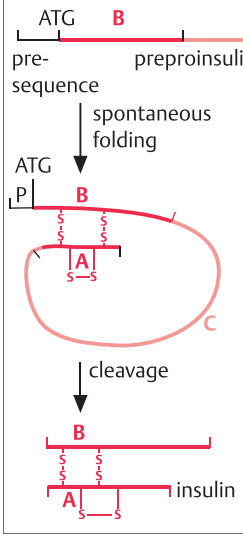
Tertiary structure

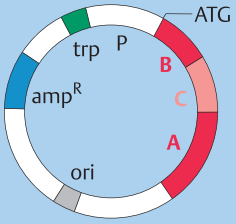


brown:
A chain;
blue: B chain;
yellow:
cystine bridges;
red: exchange of ala^{30}
(pig) against thr^{30} (man)

porcine insulin, x-ray analysis at 0.18 nm resolution (9INS)

Biosynthesis and expression plasmid





ori — origin of replication
amp^R — resistence marker
trp — promoter
P — presequence with tryptophan synthase as fusion protein, with ATG start codon (met) for BrCN cleavage

Insulins of different origin (only C-terminal end of B chain)

insulin lispro with enhanced bioavailability	recombinant human insulin from *E. coli* K12	human insulin	porcine insulin	bovine insulin*
glu	glu	glu	glu	glu
arg	arg	arg^{22}	arg^{22}	arg
gly	gly	gly	gly	gly
phe	phe	phe	phe	phe
phe	phe	phe	phe	phe
tyr	tyr	tyr	tyr	tyr
thr	thr	thr	thr	thr
lys^{28}	pro^{28}	pro^{28}	pro	pro
lys	lys	lys	lys	lys
thr^{30}	thr^{30}	thr^{30}	ala^{30}	ala^{30}

trypsin: C-terminal octapeptide is exchanged

carboxypeptidase Y: C-terminal ala is exchanged for thr

thr^{30} ← ala^{30}

* differs in another two amino acids in the A chain

Manufacture of human insulin with *E. coli* K12

bioreactor	recovery	processing	
40 m^3, 30 hrs at 37 °C, C- and N-source	isolation of inclusion bodies, removal of cell fragments	1. CNBr cleavage of fusion protein at met site 2. oxidative sulfitolysis 3. reduction 4. cystine bridges formed (O_2, pH 10.6) 5. trypsin/carboxypeptidase B: removal of C-peptide	ca. 2.5 g human insulin/ m^3 in 30 hrs

Growth hormone and other hormones

General. Growth hormone (GH, somatotropin) is a widely used hormone produced by genetic engineering techniques. It is synthesized in the anterior pituitary gland and modulates a wide range of metabolic functions. Two mechanisms are of particular importance: if food intake is high, GH inhibits fat synthesis, thus channeling energy into protein biosynthesis, e. g., in the mammary gland or muscle tissue. As a result, GH has an anabolic effect, leading to enhanced protein and reduced fat formation, and GH enhances milk production in lactating animals. A second effect of GH is mediated by the insulin-like growth factor IGF-1, which is formed in the liver and induces cell division in most tissues, thus mediating the growth-promoting activity of GH.

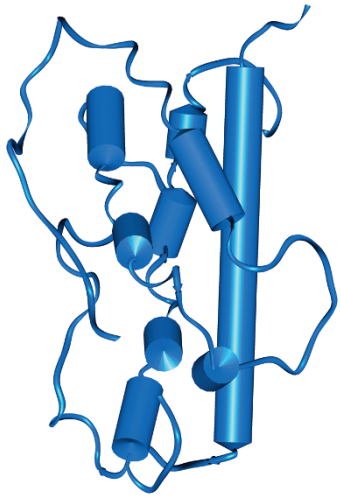
Human growth hormone (hGH) is a polypeptide composed of 191 amino acids which contains 2 disulfide bridges. Therapeutic parenteral application is widespread in children with growth retardation in which the delay in longitudinal growth is explained by an insufficient endogenous GH secretion (0.1 % freqency). In most cases, in which growth retardation is due to GH receptor defects, however, supply of exogenous GH is ineffective. Overproduction of GH may lead to excessive growth in children and to acromegaly (an excessive growth of fingers, toes, ears, and the nose) in adults. hGH is used as an anabolic drug in bodybuilding. Global sales are ca. 3 bill. US-$.

Animal somatotropin. Bovine growth hormone (bGH) differs from hGH by 67 amino acids and is also produced by genetic engineering. Since 1990, it has been registered in the USA and some other countries (but not in the EU, Canada or Japan) for increased milk production in cows, though this application is still controversial. The effect is due to enhanced energy supply to the udder. In pigs, porcine GH (pGH) is used to improve the fattening performance due to enhanced protein and reduced fat synthesis. It is predominantly used for fattening those pig breeds that produce high amounts of fat at the expense of protein. Although hGH is efficient in other mammals, both bGH and pGH do not show anabolic effects in humans. Thus, the risk of hormonally active residues in the food chain from the consumption of GH-treated milk or pork is nil, especially since these proteins are taken up via the gastrointestinal tract, in which proteins are degraded. Even if parenterally administered, GH is not stored in tissues. In aquafarming, transgenic salmon have been raised by cloning salmon GH (sGH) behind the highly active promoter of an antifreeze protein. They grow 3–10-fold larger than untreated controls. It is claimed that, due to the closed aquaculture and manipulation of the chromosomal DNA leading to infertility, the commercial production of these fish does not pose any ecological risk. The FDA might soon approve production (2014).

Fermentation and recovery. Before the advent of genetic engineering techniques, GH was obtained by extraction from pituitary glands and thus was in very limited supply and poor quality. Since 1984, it has been produced by genetic engineering, mostly through recombinant *E. coli* strains. The cloning of hGH proved quite difficult in the beginning, since the hormone and its mRNA occur in only minute quantities in the pituitary glands, and cloning experiments must make do with one single restriction site in the cDNA of hGH. The problem was solved by total synthesis of the 5'-part of the cDNA and its combination with the 3'-fragment to a functional open reading frame (ORF) that could be expressed in *E. coli*. In present industrial processes, natural hGH is formed, to circumvent any immunological risks during continuous application. Purification of the recombinant hormone is carried out via a sequence of chromatographic steps (→106).

Other recombinant hormones. Numerous other hormones have been cloned and are presently being investigated in more or less advanced tests for registration. Some are used in therapy, such as teriparatide, a recombinant fragment of human parathyroid hormone, for the treatment of osteoporosis, or exenatide (Byetta®), a glucagon-like hormone which stimulates insulin release. Recombinant human follicle-stimulating hormone (FSH) is used in some cases of human infertility. The heteromeric glycoprotein is produced in CHO cells (→98). For animal reproduction, naturally occuring analogs are preferentially used (e. g., horse serum gonadotropin PMSG from the blood of pregnant mares).

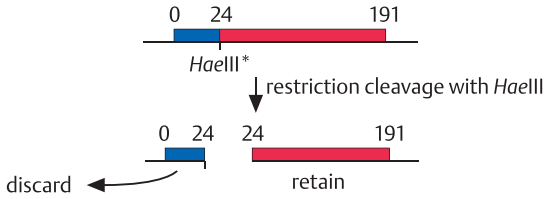
Somatotropin



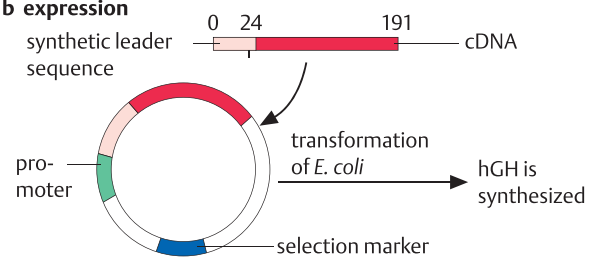
human somatotropin
at 0.25 nm resolution (1HGU)

Cloning of human somatotropin

a preparation of the cDNA fragment for hGH



b expression

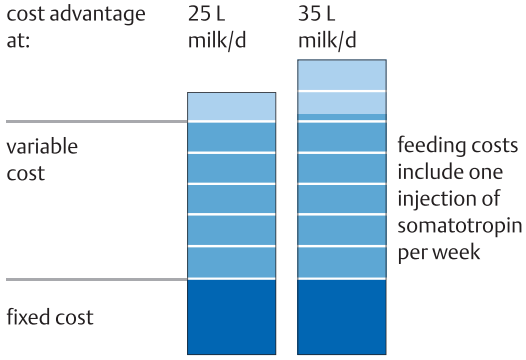


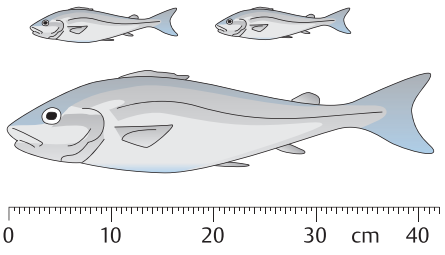
*HaeIII = restriction endonuclease from *Haemophilus aegyptus*

Fermentation and recovery

inoculation culture	bioreactor	recovery	
E. coli K12 with plasmid containing hGH gene	*E. coli* nutrient medium, 37 °C	cell lysate, chromatography, enzyme treatment resulting in charge difference, precipitation, gel and ion chromatography	hGH (example Norditropin)

Somatotropin and milk production





transgenic Atlantic salmon for aquaculture, containing a multicopy plasmid coding for sGH (upper part: controls)

Registered recombinant hormones (2013, selection)

hormone	structure	function	indication	manufacturer
glucagon	1 peptide chain, 29 amino acids	regulates gluconeogenesis and starch degradation	hypoglycemia	Novo-Nordisk
exenatide	1 peptide chain, 39 amino acids	stimulates insulin release	diabetes type II	Lilly
follicle-stimulating hormone (FSH)	2 peptide chains, 92 and 111 amino acids, glycosylated	promotes growth of egg cells, spermatogenesis	unfertility, in-vitro fertilization ivF	Merck Serono, MSD
teriparatide, parathyroid hormone fragment	1 peptide chain, 34 amino acids	increases Ca^{2+} concentration in blood plasma	osteoporosis	Lilly
thyroid-stimulating hormone (TSH)	2 peptide chains, 92 and 118 amino acids, glycosylated	stimulates formation of thyroxin	thyroid cancer	Genzyme

Hemoglobin, serum albumen, and lactoferrin

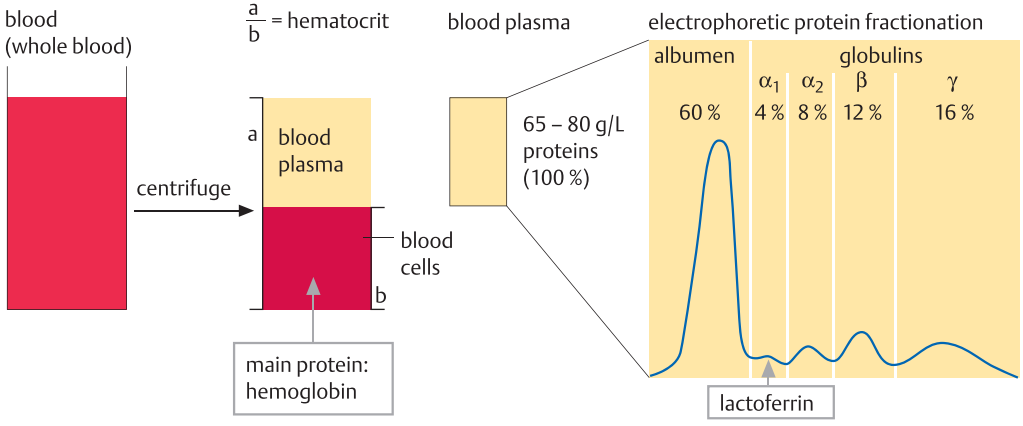
General. Blood is composed of cells suspended in serum (blood plasma). In multicellular organisms, it is the most important medium for the transport of metabolites, pH buffering, regulation of body temperature and water balance, and defense against pathogens. In humans, ca. 20 % of all genes code for blood proteins. Hemoglobin in the erythrocytes effects the transport of O_2 to some 10^{13} cells which make up our body. Sparingly water-soluble compounds are often transported in blood after binding to serum albumen. Blood proteins can be prepared by fractionating blood serum, but many blood proteins have already been prepared using genetic engineering techniques. Examples are α_1-antitrypsin (→232), an elastase inhibitor contained in blood which protects lung tissue from proteolytic degradation, and lactoferrin, an antibacterial protein in milk. The propagation of blood cells involved in defense against pathogens, as well as the formation of antibodies from B cells, is regulated by cytokines (→80). Hormones modulate numerous cell functions with high selectivity, as growth factors modulate the growth of selected types of cells. The viscosity of blood is regulated by a complex protein cascade, which prevents the formation of aggregated blood platelets through the formation of anticoagulants (→230), but is also capable of forming a clot via fibrin formation if a blood vessel has been injured. Malfunctions in this delicate balance may lead to numerous diseases. With the advent of genetic engineering, it became possible to produce proteins involved in these complex processes in substantial quantities, both for medical research and for therapeutic use.

Hemoglobin is the main protein of the erythrocyte. It is a nonglycosylated tetramer $\alpha_2\beta_2$ of M_R 64 kDa, composed of a pair of 2 identical ab subunits carrying 4 heme groups. Via allosteric regulation, binding of one O_2 increases the affinity of the other heme groups for oxygen. Hemoglobin is used in medical therapy after drastic blood loss for the transfusion of whole blood or erythrocyte concentrates obtained from blood donors – a treatment not without risks, since immunological side reactions occur and the donor blood may be contaminated with viruses. As a result, human hemoglobin has been cloned and expressed in *E. coli*, *S. cerevisiae*, or transgenic pigs. Al-though it has been obtained in high purity, the isolated protein is nephrotoxic and unstable outside the erythrocyte cell: it easily decomposes into αβ dimers, which are rapidly degraded by proteolysis. Protein engineering and microencapsulation are currently being tested for removing these disadvantages.

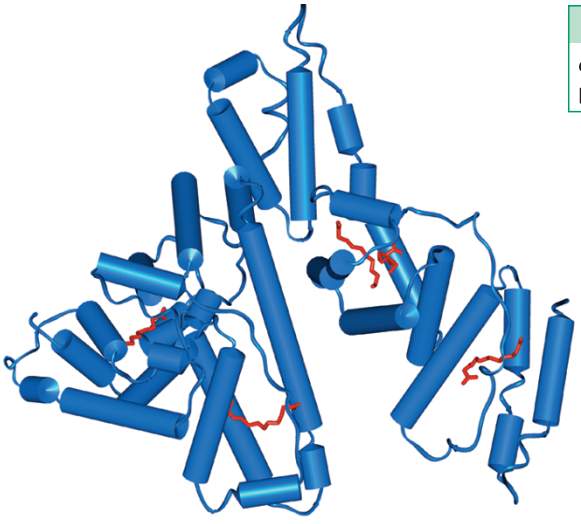
Serum albumen (hSA). is a nonglycosylated protein of M_R 69 kDa, formed as preproalbumen in the liver; it constitutes ca. 60 % of all serum proteins and thus has a significant influence on the osmotic pressure of blood. It binds and transports compounds of limited water solubility, e. g., lipids. In medicine, it is mainly used as a plasma expander in the treatment of shock after major loss of blood. It can be obtained by fractionation of donor blood, by a series of precipitation and chromatography steps. Viruses and other pathogens are removed by controlled heating of the sterile product at 60 °C for several hours. However, again and again, blood transfusions have resulted in severe infections. Recombinant human serum albumen (rHSA) has been expressed in *Bacillus subtilis, E. coli*, and *Saccharomyces cerevisiae*, and also in transgenic plants and in the milk of transgenic goats (→272). rHSA expressed in *Saccharomyces cerevisiae* (albucult®) is available from Novozymes as a reagent for animal cell culture (→100) and for the formulation of pharmaceutical preparations. Due to the manufacturing process, it is free from contamination by animal proteins.

Lactoferrin (M_R 77 kDa) is a protein with antibacterial and anti-inflammatory functions, whose activity is probably due to the high affinity with which it binds Fe^{3+} ions. The milk of lactating animals contains ca. 100 mg L^{-1} lactoferrin. In one series of experiments, it has been cloned behind the bovine α_{S1}-casein promoter, and a transgenic herd of cows or goats producing up to 10 g L^{-1} lactoferrin was successfully established (→272). The expression of this protein in transgenic tobacco or rice plants has also been reported. The recombinant product has obtained orphan drug status (GRAS) from both the FDA and EMEA, but attempts to develop it into a drug for an orally administered treatment of ulcers or non-small cell lung cancer have failed so far.

Blood proteins



Albumen by fractionation of donor blood



human serum albumen (1 HSA) at 0.24 nm resolution;
5 bound molecules of myristic acid (red)

donor blood: human plasma

↓

filtration, desalting
desalted blood plasma

↓

DEAE Sepharose
pH 5.2: immunoglobulins

↓

DEAE Sepharose
pH 4.0: glycoproteins, ceruloplasmin

DEAE Sepharose
pH 4.4: albumen fraction

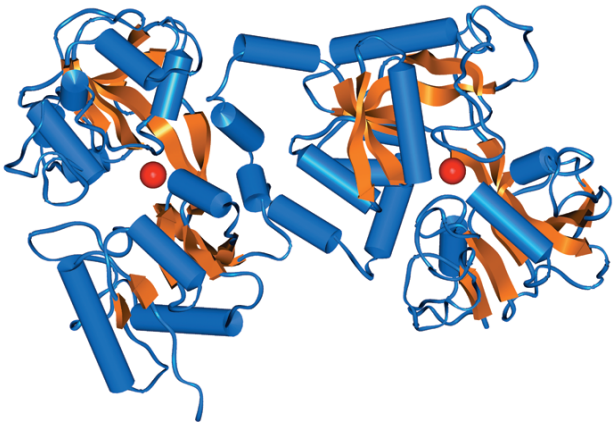
↓

CM Sephadex
ion gradient, pH 4.2 – 8.1: various proteins, pH 5.5: albumen

↓

ultrafiltration, Sephacryl
pure albumen

Lactoferrin



human lactoferrin (1BJ5) at 0.24 nm resolution.
Two chelated Fe ions in red.

transgenic cow
with human lactoferrin cDNA, cloned to regulation elements of the αS1-casein promoter of beef

↓

extraction
addition of 0.4 M NaCl to milk, extraction with S-Sepharose

↓

quality control
peptide mapping, MALDI-TOF mass spectrometry, 2D gel electrophoresis, glycosylation pattern

Blood clotting agents

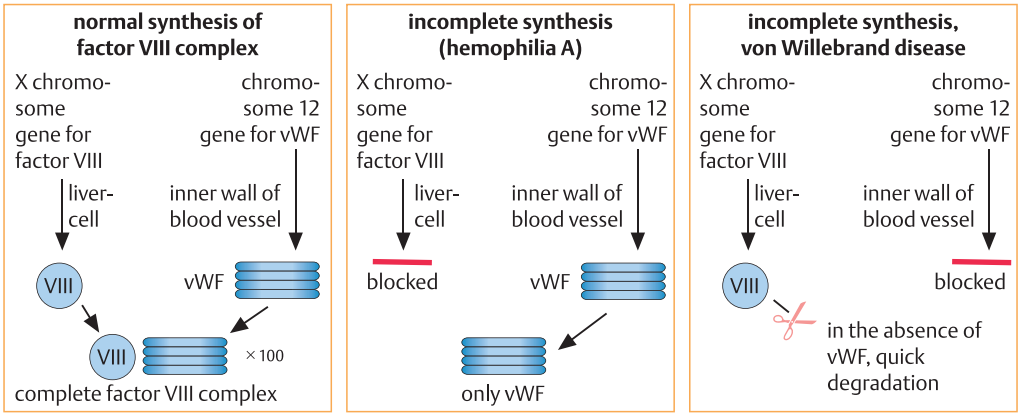
General. Once blood vessels have been damaged from within or outside, a blood clot forms quickly, preventing further bleeding. This process, called hemostasis, is regulated by a complex cascade of interactive steps, involving zymogen activation, proteolysis, and inhibition of proteolysis, which prevent blood clotting in the healthy organism. In an initial step to blood clotting, soluble fibrinogen is hydrolyzed by a protease, resulting in the formation of a soft clot (via an altered charge distribution) which is transformed into a hard clot through the action of transglutaminase (factor XIIIa), which forms additional amide ("isopeptide") bonds (→192). The serine protease that hydrolyzes fibrinogen is called thrombin. It is liberated from a precursor protein, prothrombin, by the action of factor X_a, another protease, whose activity is modulated by the factor VIII complex. Most bleeders suffer from genetic aberrations that influence the synthesis or function of factor VIII or X. **Hemophilia.** Descriptions of hemophilia were found on ancient Egyptian clay tablets. Today, 3 major diseases are distinguished: hemophilia A, hemophilia B, and von Willebrand disease. Hemophilia A occurs with a frequency of 1:5,000 only in males. It is the consequence of a defect in the biosynthesis of the factor VIII complex: if < 1 % of the normal amount of this complex is formed, spontaneous bleeding may occur, and life expectancy is low. An unusual inversion on intron F8A, preventing biosynthesis of the gene product in the liver, often seems to be the cause of this defect. The factor VIII gene is localized on the X chromosome. Factor VIII is a glycoprotein (M_R ca. 300 kDa), composed of 2,332 amino acids forming a single peptide chain. The sugar content of the glycoprotein is ca. 35 %, with 25 putative glycosylation sites. A structural model has been obtained by electron crystallography, and an x-ray structure of the deglycosylated protein is available. Biosynthesis of factor VIII proceeds by splicing a DNA segment ca. 186 kbp long and containing 26 exons. During posttranslational glycosylation, the B domain is heavily glycosylated; it is later removed during activation by thrombin. von Willebrand disease is due to an erroneous biosynthesis of the von Willebrand factor (vWF) on the inner wall of blood vessels. The vWF gene is located on chromosome 12. As a result, the disease occurs with equal frequency in men and women (1:1,000). The vWF protein is even larger than factor VIII, and just as glycosylated. Up to 100 monomers associate into large multimers, and each monomer binds one molecule of factor VIII, yielding the factor VIII complex (VIII:vWF). It accelerates over 100-fold the aggregation of blood platelets through the activation of the factor X/factor IX_a system. Hemophilia B occurs with a frequency of 1:25,000 and mostly among men. It is caused by an erroneous synthesis of factor IX, a glycoprotein of M_R ca. 55 kDa. Factor IX, like the factor VIII complex, participates in the activation of factor X to the active factor X_a. Its gene is localized on the X chromosome (Xq27) along a 34 kbp stretch.

Cloning. The cloning of factor VIII was achieved nearly simultaneously by Genentech and Genetics Institute, the low occurrence of an mRNA transcript (only 10^{-5} of the total mRNA of liver) constituting the major challenge. A complete cDNA transcript was eventually obtained from a lymphoma cell line through genome walking (→70), cloned into a vector containing elements of SV40 and adenovirus promoters, and functionally expressed both in CHO- and BHK-cell lines (→100).

Manufacture. Since ca. 1964, factors VIII, IX, and vWF have been isolated in pure form from donor blood by cryoprecipitation and fractional immunochromatography (→106, 226). After freeze drying, it is used for parenteral therapy. Since blood from several thousand donors per annum is required to supply a single hemophiliac A with factor VIII, the risk of viral infection is very high; estimates run as high as > 60 %. Against this background, the successful manufacture of recombinant factors VIII and IX since 1992 constitutes a major breakthrough. Due to the high degree of glycosylation (→262), animal cell cultures of CHO cells are used as hosts (→100). The yields are low, amounting to some mg of product per L of cell culture. As a result, the product is very expensive, and the market value was estimated at > 2 billion USD (2006), in spite of a relatively small number of patients. Present developments aim for a "long-acting factor VIII" through protein engineering (→198), modification of sugar chains (→262) or binding of polyethylene glycol (PEGylation).

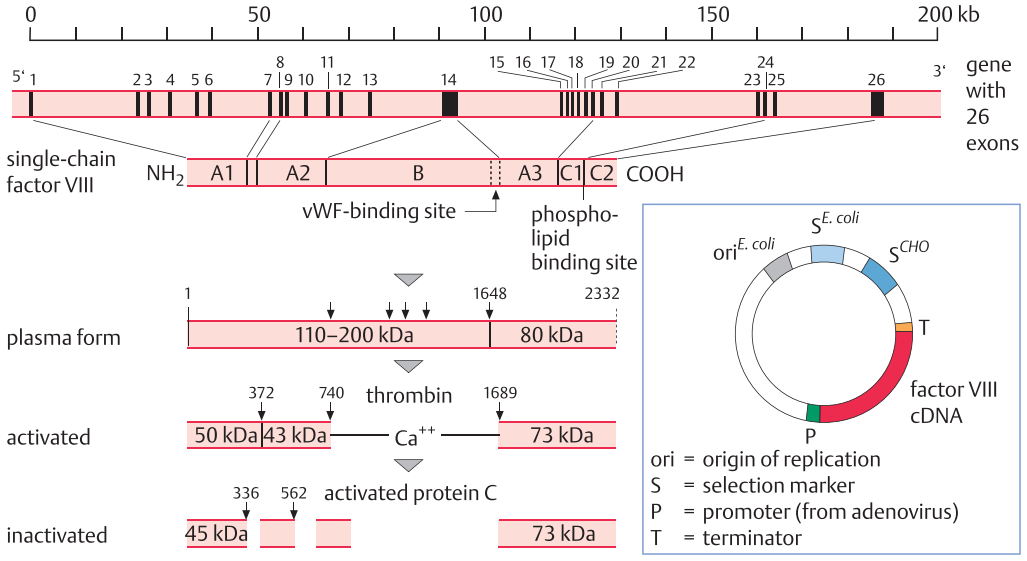
Frequent hemophiliac diseases

	hemophilia A	von Willebrand disease	hemophilia B
inheritance	1:5,000, males only	1:1,000, men and women	1:25,000, mostly men
clinical symptoms	bleeding at joints and muscles, brain hemorrhage	nose bleeds, strong menstrual bleeding, prolonged bleeding of wounds	spontaneous joint bleeding in child-hood
locus	Xq 28	12p12	Xq 27



The factor VIII complex is composed of factor VIII and the von Willebrand factor (vWF).

Factor VIII: gene structure and expression vector



Manufacture of factor VIII

conventional procedure

500 blood donors
100 L plasma,
1 kg cryoprecipitate

→ **purification**
precipitation, multi-stage immunochromatography

→ von Willebrand factor → 1 mg factor VIII

recombinant production

recombinant CHO cells

→ **cell culture reactor**
200 h , 35 °C,
complex medium

→ **recovery**
precipitation, multi-stage immunochromatography

→ some mg of factor VIII/L cell culture

Anticoagulants and thrombolytic agents

General. The sudden formation of a thrombus (vein occlusion, heart attack, stroke) is among the most frequent causes of death in industrialized nations. In the USA, ca. 2.5 million people die each year from the consequences of cardiovascular diseases. Anticoagulants prevent the formation of primary thrombi (e. g., after surgery). Thrombolytic agents dissolve thrombi by proteolysis. Important anticoagulants are heparin, coumarin derivatives, and thrombin inhibitors such as hirudin or human antithrombin-III (AT-III), produced by genetic engineering techniques. Thrombolytic agents are bacterial streptokinase and the recombinant preparations urokinase and tissue plasminogen activators (tPA, reteplase, TMK-tPA).

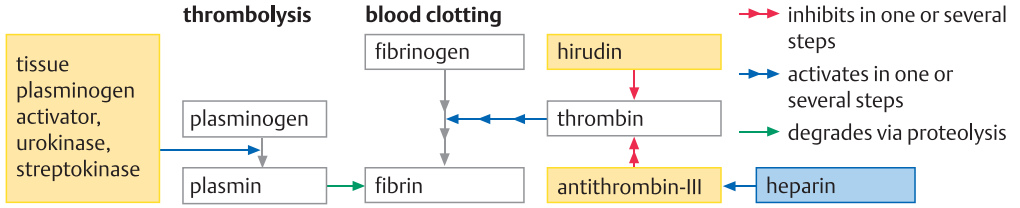
Heparin is a polydisperse sulfated glucosaminoglucan of M_R 3–60 kDa. It can be obtained by extraction from porcine intestine or bovine lung. It is formed by mast cells and secreted into blood plasma, where it activates AT-III, which in turn inhibits the formation of fibrin by binding to thrombin.

Hirudin is a thrombin inhibitor, originally isolated from the saliva of leeches. It has been expressed in *E. coli*, *Hansenula polymorpha*, and other host organisms. As a result, it can be produced by fermentation (e. g., Lepirudin™). Like AT-III, it inhibits the formation of fibrin by binding to thrombin.

Tissue plasminogen activator, tPA. The degradation of fibrin, as observed during wound healing, is to a considerable extent catalyzed by the serine protease plasmin. However, this reaction proceeds only if plasmin has been formed from its zymogen plasminogen through the action of tPA, a serine protease. tPA thus is a thrombolytic agent. It has a molar mass of 72 kDa. It hydrolyzes, with high selectivity, the Arg^{561}-Val^{562} peptide bond of plasminogen. tPA forms five domains, whose functions were derived from their homologies to other proteins. The two "kringle domains" bind to the substrate fibrin, and the protease domain contains the functional center of the enzyme. Human tPA was first cloned in 1982 and has been commercially available since 1986. Its 8 disulfide bonds and 3 sugar side chains, originally believed essential for substrate binding (ca. 5 % of the overall molar mass), rendered functional expression in *E. coli* impossible. Production of the recombinant enzyme is therefore carried out in CHO cells (→98), followed by a complex series of purification steps which include precipitation, ion-exchange chromatography, and immunoaffinity chromatography (→106). An engineered variant of tPA, with 4 amino acid changes and moved glycosylation sites (Tenecteplase, TNK-tPA®) (→262), has improved fibrin selectivity and a longer half-life in serum, allowing for "single-bolus" dosage instead of infusion. Non-glycosylated mutants of tPA are also active and can be functionally expressed in *E. coli*. Thus, reteplase (Repalysin®), a mutant without the kringle-1 and the epithelial growth factor-domain, is manufactured in an *E. coli* host and must be refolded from inclusion bodies (→104); it exhibits a 3–4-fold prolonged residence time in blood serum and is not allergenic. tPA has also been expressed in the milk of transgenic animals, after the host animal had been transformed with a vector where the tPA cDNA had been cloned behind a lactalbumin promoter (→272).

Other thrombolytic agents. Urokinase is a serine protease that is synthesized in the urogenital tract. It is formed as pro-urokinase in plasma and urine. Like tPA, urokinase hydrolyzes plasminogen to plasmin. Two variants of similar biological activity (M_R 54 and 30 kDa) can be purified separately, the lighter variant originates from the heavier one by autolysis. Urokinase can be prepared from urine, from human kidney cultures, or from recombinant *E. coli*. Streptokinase is a catalytically inactive protein of M_R 45 kDa that is formed by some hemolytic streptococci. Once bound to plasminogen, it induces a conformation change in plasminogen, resulting in autolytic degradation to plasmin. Streptokinase is obtained from the supernatant of streptococci cultures and purified by chromatographic procedures (→106). While the production cost of this thrombolytic agent is favorable, streptokinase bears the risk of immunological reactions. In the saliva of vampire bats, several plasminogen activators (bat-PA) have been found and cloned. The largest (desmoteplase) has 477 amino acids and a function very similar to tPA, but leads to an immune response in humans and thus has not yet been cleared for application.

Thrombolysis (much simplified)



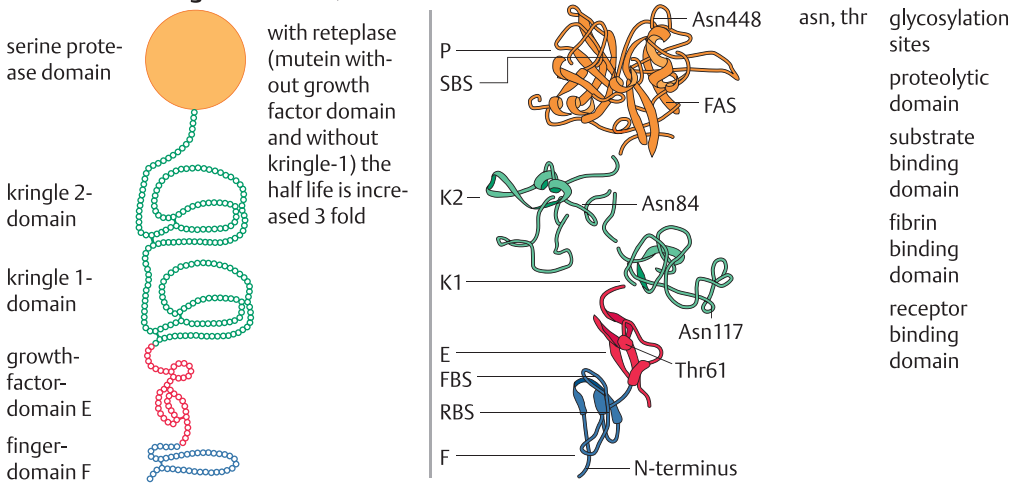
Anticoagulants (A) and thrombolytic agents (T)

	type	effect	enterprise (selection)*
heparin	A	sulfated polysaccharide binds to anti-thrombin III and inactivates thrombin	Celsus
hirudin	A	inhibits thrombin	Novartis, Schering*
antithrombin-III	A	inhibits thrombin	Genzyme
streptokinase	T	activates plasminogen	Kabi Upjohn*
urokinase	T	activates plasminogen	Grünenthal*
tPA	T	activates plasminogen	Genentech*, Boehringer Ingelheim*, Roche*

*registered

natural compound recombinant protein

Tissue Plasminogen Activator, tPA



Manufacture of tissue plasminogen activator (examples)

cell reactor	**recombinant CHO cells** vector CHO NEOSPLA	**mammalian cell reactor** > 100 m^3 fixed bed reactor, complex medium, 37 °C, 20 d	**recovery** several immunochromatographic steps	several mg/L·d
transgenic animal	**transgenic gout** vector LSAP + PA	**animal farming** 240 d milk production 3 – 4 L/d	**recovery** acid precipitation, butyl-Sepharose, immunochromatography	up to 30 g/L milk

quality control

complex quality control using RP-HPLC, SDS-PAGE, ELISA, peptide mapping, analysis of glycosylation pattern, and other methods

Enzyme inhibitors

General. Enzyme inhibitors have found a firm place in therapy. For example, the protease inhibitor aprotinin, obtained from bovine tissues, is used to treat certain types of pancreatitis or shock. In the future, recombinant α_1-antitrypsin (→226) might be used for treating emphysema. Although protease inhibitors of microbial origin, such as the leupeptins, pepstatin, antipain, chymostatin, and elastinal have not yet found a therapeutic use, the microbial metabolite acarbose, a glycosidase inhibitor, is a valuable antidiabetic. Oseltamivir is used to treat influenza and bird flu. Tetrahydrolipstatin, a chemical derivative of the microbial metabolite lipstatin, has found broad use in the treatment of obesity, since it inhibits human pancreatic lipase.

Aprotinin is a polypeptide built from 48 amino acids (M_R 6511), which inhibits various proteases such as trypsin, chymotrypsin, and plasmin (pancreatic trypsin inhibitor, PTI). The inhibition constant K_i for trypsin is ca. 10^{-11} M. Aprotinin (Trasylol®) is used for treating pancreatitis and, under strictly controlled conditions, strong bleeding, shock, and organ transplantation as well. Still another application is in fermentations using mammalian cells, where aprotinin may prevent proteolysis of the recombinant proteins secreted into the medium. Aprotinin is isolated by extracting bovine pancreas or lungs, followed by chromatographic purification. It is not glycosylated and thus can also be functionally expressed in *Escherichia coli* host cells.

α1-Antitrypsin (αAT). (→226) This large glycoprotein (M_R 54 kDa) is coded on chromosome 14 (14q32). It is synthesized in the liver and circulates in the blood serum at a concentration of ca. 2 g L^{-1}, corresponding to > 90 % of the α_1-globulin fraction. It inhibits elastase, a protease secreted from the neutrophilic granulocytes of the immune system, preventing the proteolysis of lung tissue, which is largely composed of the protein elastin. In a genetic defect that occurs predominantly in Northern Europe, αAT is mutated at position 53 (lys^{53}→glu, "Z-type"). In this mutation, the secretion of αAT by liver cells is greatly reduced, resulting in a serum level of just 15 % of the normal value. As a consequence, elastase starts to hydrolyze lung tissue, leading to life-threatening emphysema and to usually fatal adult respiratory distress syndrome. Smokers with Z-type αAT are especially endangered, since components of tobacco smoke oxidize met^{358}, which is essential for elastase inhibition. By intravenous application of the inhibitor (ca. 200 g/patient/year), the progress of disease can be retarded. αAT is predominantly obtained by plasma fractionation of donor blood (→226). The inhibitor is also prepared in recombinant form. Since it requires a complex glycosylation pattern for its biological function, expression in *Saccharomyces cerevisiae* (→262) led to a better product than *E. coli*. The most economically attractive alternative seems to be expression of the recombinant glycoprotein as a fusion product with β-lactoglobulin, secreted into the milk of transgenic sheep (→272).

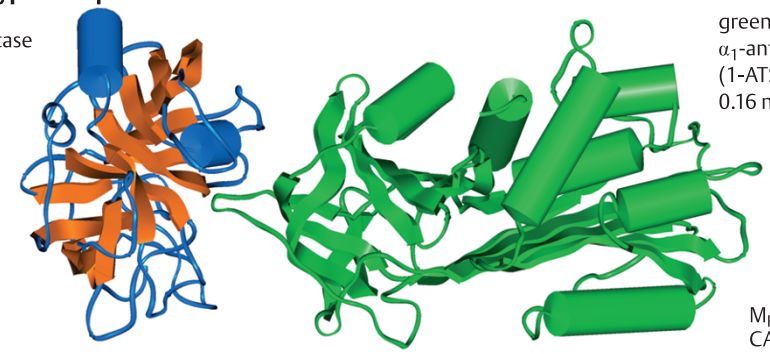
Acarbose (Glucobay®), a pseudotetrasaccharide produced by the Actinomyces strain *Actinoplanes utahensis*, is a competitive inhibitor of invertase, maltase, α- and β-amylase, and various glucosidases. It reduces the glucose content in the gastrointestinal tract and thus is used as an oral antidiabetic and anti-adiposis drug, with annual sales on the order of 300 million US$. It is manufactured by microbial fermentation. Several genes involved in its biosynthesis have been cloned.

Oseltamivir (Tamiflu®) is a prodrug for a synthetic inhibitor with virostatic action. After oral ingestion, it is hydrolyzed by liver esterases to the active acid form which inhibits the enzyme neuraminidase, localized on the surface of adenoviruses. As a consequence, the virus becomes unable to hydrolyze sialoglycoproteins which cover the surface of the host cell, and the infection becomes less virulent. In 2007, the WHO recommended the use of the inhibitor as a preventive measure against the H5N1 bird flu pandemic (→250).

Lipstatin is a lipophilic ester with a mid-chain β-lactone ring and an *N*-formyl-L-leucine side chain, produced by *Streptomyces toxytricinii*. Through catalytic hydrogenation, it is transformed into tetrahydrolipstatin (Xenical®). Both compounds bind covalently to the serine residue in the active site of many lipases. After oral ingestion, it blocks pancreatic lipase, leading to an inhibition of triglyceride hydrolysis without reducing the uptake of free fatty acids. Tetrahydrolipstatin is used as a treatment for obesity; its patent protection expired in 2009.

α_1-Antitrypsin – a protease inhibitor

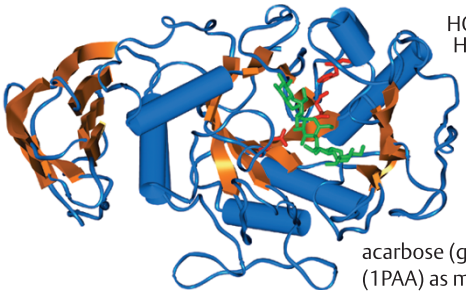
human elastase (1-EGS) at 0.16 nm resolution

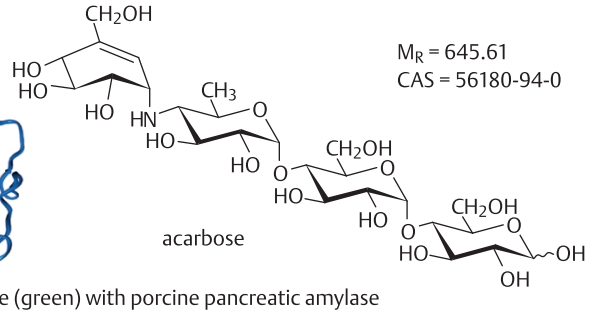


green: human α_1-antitrypsin (1-ATS) at 0.16 nm resolution

M_R = 54 kDa
CAS = 9041-92-3

Acarbose – an α-glucosidase inhibitor



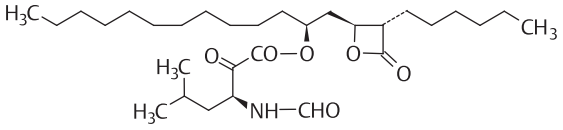


acarbose

M_R = 645.61
CAS = 56180-94-0

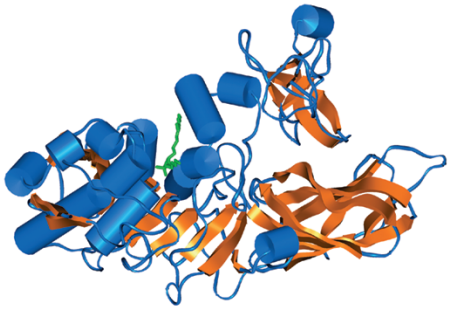
acarbose (green) with porcine pancreatic amylase (1PAA) as model target at 0.23 nm resolution

Tetrahydrolipstatin (Xenical®) – a lipase inhibitor



M_R = 495.74
CAS = 96829-58-2

tetrahydrolipstatin (green) with human pancreatic lipase (1 HUP) at 0.23 nm resolution



Some processes and yields

α_1-antitrypsin	acarbose	tetrahydrolipstatin
transgenic founder sheep Tracy	**preculture** high-performance mutants of *Actinoplanes utahensis*	**preculture** high-performance mutants of *Streptomyces toxytricinii*
milk	**bioreactor** several m^3 starch or maltose, 5 – 6 d at 28 °C	**bioreactor** several m^3 starch or dextrins, soy meal, 124 h at 28 °C
recovery precipitation of casein, chromatography	**recovery** filtration, ion-exchange chromatography	**recovery** filtration, extraction with ethyl acetate, reverse-phase chromatography
10 mg pure αAT per L milk	several grams/L	several grams/L

Interferons

General. Interferons (IFN) are proteins synthesized in various cells of the immune system (→80). They serve as messengers that initiate an immune response after binding to IFN receptors. Higher animals form several types: IFN-α, IFN-β, IFN-γ, and IFN-ω, which participate in the regulation of ca. 20–30 genes and exhibit a broad spectrum of immunomodulating, antiviral, and antiproliferative properties. IFN-α and IFN-β are stable at pH 5 and bind to the same receptor (type I-IFN), whereas IFN-γ is acid labile and binds to type II-IFN receptors.

Properties and applications. α-Interferons (IFN-α) are formed by leukocytes as a family of >20 nonallelic genes that exhibit high sequence homology. Their protein chains consist of 165–166 amino acids with a molar mass of ca. 16 kDa, which can increase up to 26 kDa by glycosylation. Up until now, the most important clinical applications of IFN-α are in the treatment of hepatitis B and C and of cancers such as bladder tumors, melanoma, leukemia, and lymphoma. The market value (world) is ca. 1.4 billion US$ y^{-1}. β-Interferon (IFN-β) is synthesized by fibroblasts. Its protein chain consists of 166 amino acids; but due to glycosylation, its molar mass is ca. 20 kDa. Its major clinical use is in the treatment of multiple sclerosis. γ-Interferon (IFN-γ), sometimes termed "immuno interferon", is formed by activated T lymphocytes and in turn activates lymphocytes. Its peptide chain, made up of 143 amino acids, can be glycosylated to a various extents, resulting in molar masses between 15 and 25 kDa. IFN- γ1b is registered for the therapy of chronic granulomatosis and osteoporosis. Further therapies based on the use of interferons are in various stages of clinical testing, e. g., the therapy of several cancers (IFN-α, -β and γ), of autoimmune diseases and viral infections (IFN-α, -β and -ω), and of rheumatoid arthritis, idiopathic pulmonary fibrosis and asthma (IFN-γ). Interferons are usually administered by intramuscular injection, but viral respiratory diseases have also been treated by intranasal administration. The world market for all variants of interferons is an estimated 10 bill US$ (2014).

Cloning and expression. Although the therapeutic potential of the interferons was quickly recognized, their classical preparation by fractionation of donor blood prevented any large-scale clinical investigations. Only the advent of genetic engineering led from 1986 onward to the industrial manufacture of pure interferons as clinical preparations. The cloning of these proteins, which are formed only in very low quantities, was first successful for IFN-α in 1982, with the following methods: 1) isolation of leukocyte mRNA and reverse transcription into cDNA, 2) expression in *E. coli* and hybridization with mRNA coding for IFN-α, 3) translation of the hybridizing mRNA by cell-free protein synthesis and testing the antiviral properties of the resulting protein. Since the glycosylation pattern of the recombinant interferons does not exert a major influence on their clinical effects, *E. coli* is often used as the host organism. However, functional interferons were also expressed in other host cells, such as *Saccharomyces cerevisia*, *Pichia pastoris*, and mammalian cells.

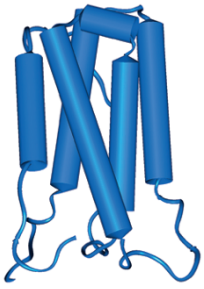
Manufacture and recovery. In ca. 1978, the industrial manufacture of IFN-α began based on human lymphoblastoma cell lines infected with Sendai virus (*Nawalma* cells). This procedure, however, led to the formation of at least 8 IFN-α isoforms. Today, interferons are produced using recombinant *E. coli* cells or, if glyosylation is important, as in the case of IFN-ß1a, by recombinant CHO cells (→98). If *E. coli* cells are used, high yields of inclusion bodies can be obtained in high-cell-density fermentations (→92). The cost-determining step, however, is in the downstream processing, which includes refolding of inclusion bodies and chromatography (→106). As an example, Roche prepares interferon-α2a (Roferon A®) using recombinant *E. coli* K12 cells. After harvesting, the cells are destroyed by deep freezing, rendering the cell pellets containing the recombinant inclusion bodies stable for storage. After stirring in buffer, the microbial debris is removed by centrifugation, and the protein mixture is purified in several chromatographic steps. Quality control of the products, in particular their correct folding, is precise and costly. Since 2001, various interferons which have been modified by polyethylene glycol have been authorized both by the FDA and EMEA ("PEGylated interferons", at present IFN- α2a, IFN- α2b und IFN-ß1a). Through PEGylation, the residue time in the body is prolonged to about 1 week.

Registered interferons (selection)

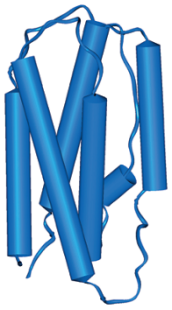
Product	trade name	manufacturer	registered since	indications
IFN beta	Fiblaferon	Rentschler	1983	virus infections (e.g., SARS)
IFN alpha-2a	Roferon A	Roche	1987	cancer therapy
IFN gamma	Polyferon	Rentschler	1989	rheumatoid arthritis
IFN gamma-1b	Imukin	Boehringer, Ingelheim	1992	chronic granulomatosis
IFN beta-1b	Betaferon	Bayer Pharma	1995	multiple sclerosis
IFN beta-1a	Avonex	Biogen	1997	multiple sclerosis
IFN beta-1a	Rebif	Merck Serono	1998	multiple sclerosis
IFN alphacon 1	Inferax	Yamanouchi	1999	hepatitis C
IFN alpha-2b	Intron A	MSD	2000	hepatitis B/C, malign melanoma
IFN alpha-2b, PEGylated	PegIntron	MSD	2002	hepatitis C
IFN alpha-2a, PEGylated	Pegasys	Roche	2002	hepatitis B/C

PEG = polyethylene glycol

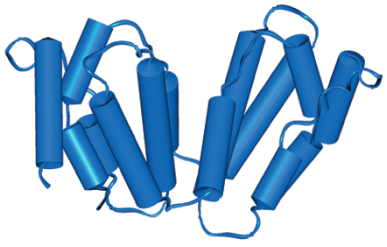
IFN-β1a is produced with CHO cells, all other interferons with recombinant *E. coli* cells



human IFN-α (1ITF), NMR data

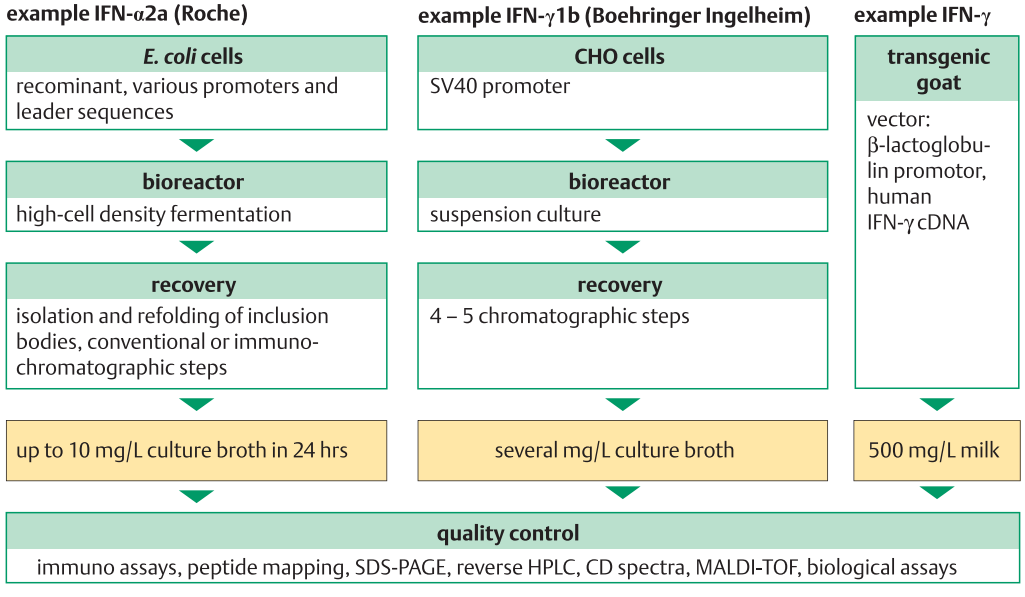


bovine IFN-β (1A41), 0.28 nm resolution



human IFN-γ, (1DC9), 0.3 nm resolution

Manufacture



Interleukins and "anti-interleukins"

General. Interleukins (IL) are proteins formed by various cell types of the immune system (→80). They are often called "the hormones of the immune system," since they modulate the activity of other cells in the immune system by binding to specific receptors. In humans, more than 30 types of interleukins have been discovered (IL-1 – IL-31)(2013). IL-2 has already been registered for therapeutic use; others such as IL-21 are in clinical trials, but side effects are usually high, retarding clinical applications. Anti-interleukin antibodies are investigated for the treatment of some interleukin- associated diseases.

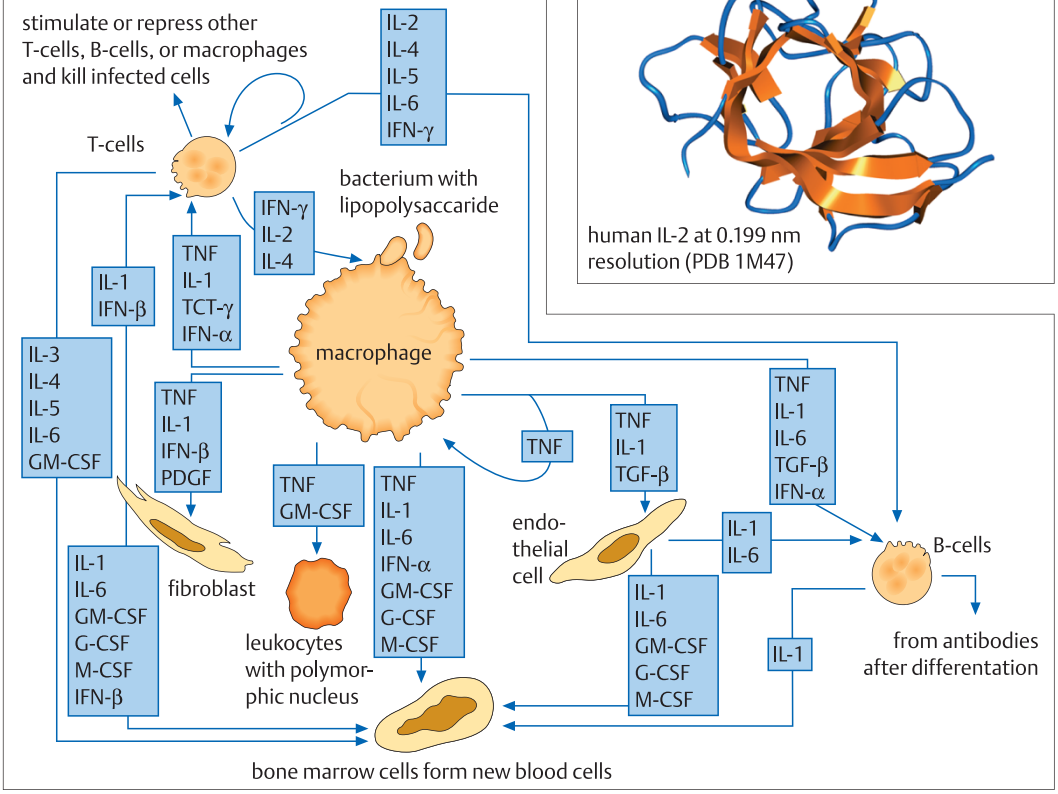
Properties and applications. Interleukin-1 (IL-1) is synthesized in two forms (IL-1α and IL-1β) by phagocytic cells (macrophages, monocytes). It is formed as a preprotein of molar mass 31 kDa, to be proteolytically processed to biologically active IL-1 of molar mass 17.5 kDa. It shows proinflammatory activity and stimulates the growth of lymphocytes, fibroblasts, hematopoetic cells, and thymocytes. IL-1 receptors have been identified on the surfaces of T-cells, fibroblasts (type I), and B-lymphocytes (type II). Upon phagocytosis and proteolysis of an antigen, macrophages are thought to secrete IL-1 and thus initiate the increased formation ("expansion") of immunodefensive cells against this type of antigen. During the subsequent amplification of the immune defense, interleukin-2 (Il-2, T-cell growth factor) plays an important role. This best-studied interleukin is formed by antigen-activated T-cells and stimulates the growth and differentiation of T- and B-lymphocytes. The immune cascade is further enhanced by the presentation of IL-2 receptors on the surface of T-cells. Since NK-cells ("killer cells") and monocytes present IL-2 receptors constitutively, their activity is also potentiated by IL-2. The crystal structure of the hydrophilic glycoprotein composed of 133 amino acids (M_R 15.5 kDa) has been solved. IL-2 is a registered drug for the treatment of kidney carcinoma. Interleukin-3 (IL-3) is a 133 amino acid glycoprotein. It is synthesized by activated T-lymphocytes and induces the differentiation of bone marrow stem cells to mature leukocytes. Since it is also involved in the differentiation of other cell types of the immune system, it is often called a "multipotent growth factor". Interleukin-4 (IL-4), a glycoprotein of M_R 20 kDa, not only stimulates the formation of B- and T-lymphocytes, like IL-1, but also the secretion of immunoglobulins IgG and IgE. In addition, it enhances the presentation of antigens by monocytes. Interleukin-6 (IL-6) has properties similar to IL-1 and IL-2, but, in contrast to them, induces the expression of several acute-phase proteins in hepatocytes; it is thought to play a role in several autoimmune diseases. Interleukin-10 (IL-10) plays an inhibiting role in the biosynthesis of other interleukins ("cytokine-synthesis inhibiting factor"). Interleukin-12 (IL-12) stimulates the synthesis of γ-interferon in T-lymphocytes and NK-cells and seems to play a pivotal role in the orchestration of immune defense. A major part of the clinical studies based on recombinant interleukins concerns cancer therapy. Further potential applications are seen in wound healing and the improvement of immune defense in AIDS patients and in immunosuppressed patients after bone marrow transplantations. The crystal structures of IL-19 and IL-22 have been solved. Their structures differs considerably from the structure of IL-2. "Anti-interleukins" are monoclonal antibodies (→242) directed against interleukins or interleukin receptors. Kineret® is an antibody registered for the treatment of rheumatoid arthritis; it acts on the IL-1 receptor. Ustekinumab (Stelara®) is directed towards IL-12 and registered for the treatment of severe plaque psoriasis. Other antibodies against interleukins for the treatment of, e. g., melanoma or asthma are under clinical investigation.

Manufacture of IL-2. Since the glycosylation pattern of this interleukin does not seem to be decisive for therapy, the protein is produced with transformed cells of *Escherichia coli*. Using medium- or high-cell density fermentations, high concentrations of inclusion bodies are obtained, which are first purified by gel permeation chromatography, then solubilized under reducing conditions. This is followed by refolding in the presence of oxidants (→104). For further purification, precipitation reactions, HPLC, and gel filtration steps are used. The activity of IL-2 is usually determined in a complex bioassay based on the ingestion of ^{3}H-labeled thymidine into Il-2 dependant T-cells of mice.

Interleukins and "anti-interleukins" (selection)

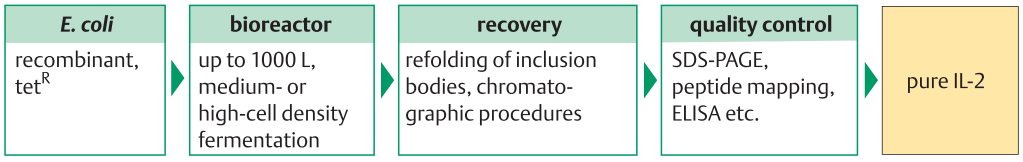
Product	Trade name	Clinical application	Producer	Status
IL-1 receptor antagonist	Kineret	rheumatoid arthritis	Amgen/Swedish Orphan	approved
IL-2	Proleukin	kidney cell carcinoma, metastatic melanoma	Novartis	approved
IL-2 PEG		HIV treatment	Chiron, stopped	
IL-3		stem cell expansion	Sandoz, stopped	
IL-4		lung cancer	Schering-Plough, stopped	
IL-6 nanobodies		Inflammatory diseases	Merck-Serono	research
IL-6 antibodies	Actemra	rheumatoid arthritis	Roche/Chugai	approved
IL-10	Tenovil	Morbus Crohn	Schering-Plough, stopped	
IL-11		marker for gastro-intestinal carcinoma		research
IL-12/IL-23 Ustekinumab-antibodies	Stelara	Psoriasis	Janssen-Cilag	approved
IL-15 superagonist protein complex		metastatic melanoma	Altor BioScience	clinical studies
IL-17 antibodies		Psoriasis	Amgen, Eli Lilly, Novartis	clinical studies
IL-21		advanced melanoma	Zymogenetics	orphan drug

Interactions in the immune system



human IL-2 at 0.199 nm resolution (PDB 1M47)

Production and downstream processing of IL-2



Erythropoietin and other growth factors

General. During the development of animal tissue culture, various factors were found to stimulate the growth of a particular cell type (colony-stimulating factors, CSF). They are usually formed in minute amounts, act by binding to surface receptors of target cells, and belong to the cytokine group (→80). Only after the advent of genetic engineering techniques did it become possible to prepare them in large amounts, to study their composition, and to explore their therapeutic potential in enhancing the growth of specific cell varieties, e. g., from skin, nerve, blood, or bones. During these early studies, erythropoietin (EPO), granulocyte CSF (G-CSF), and granulocyte-macrophage CSF (GM-CSF) were discovered to have a major medical uses: since EPO induces the formation of erythrocytes, G-CSF of neutrophilic granulocytes, and GM-CSF of eosinophilic and neutrophilic granulocytes, these proteins can be successfully employed in various kinds of anemia, in particular for dialysis patients suffering from kidney failure. In the USA alone, the number of renal anemia (dialysis) patients receiving EPO is ca. 400,000 (2012). GM-CSF is a recombinant growth factor used for the treatment of neutropenia (neutrophils in blood are reduced, e. g., after chemotherapy or dialysis; as a result, the risk of infection increases). The world market for both products is an estimated 12 billion US$ (2013) but this is decreasing as the original patent protection has expired and more biosimilars are entering the market.

Erythropoietin is a factor that stimulates the growth of erythrocytes. It is synthesized in kidney endothelial cells and in the Kupffer cells of the liver and is regulated by the partial pressure of oxygen in the blood. In hemapoetic stem cells of the bone marrow, EPO induces the loss of nuclei and the concomitant formation of hemoglobin, resulting in the formation of erythrocytes. As a result, EPO is a valuable therapeutic agent for anemic conditions, in particular for secondary anemia, which is a major consequence of blood dialysis in patients dependant on an artificial kidney. If combined with other bone marrow growth factors such as GM-CSF or G-CSF, which stimulate the growth of granulocytes, eosinophilic granulocytes, and monocytes, EPO has become an indispensable restorative drug for dialysis patients. EPO is also well known for its history in blood doping in endurance sports. The gene coding for EPO was first cloned in 1984 from human bone marrow. The glycoprotein is based on a single peptide chain of just 165 amino acids; 4 large sugar chains on three asparagine and one threonine residues of this peptide contribute ca. 40 % of its 34 kDa molar mass and are indispensable to its function. Due to the flexible sugar chains, the crystal structure of EPO remained elusive for a long time, but has recently been solved in a form bound to its receptor. Due to the need to preserve the highly specific glycosylation pattern (→262) for function, mammalian cell culture is presently the only way to manufacture EPO commercially, mostly in Chinese hamster ovary (CHO) cells (→98).

Growth factors. This large group of proteins stimulates the specific growth and differentiation of individual cell types occuring in nerves, skin, bone, connective tissue and elsewhere. Many growth factors were cloned and investigated for their clinical potential, but side effects have prohibited their use with only few exceptions such as G-CSF and GM-CSF. Growth factors are frequently used as targets for the development of recombinant antibodies directed towards the uncontrolled growth of specific cell types, e. g., in cancer treatment. Stem cell therapy (→78, 306) is an emerging and competing new concept for the expansion of specific cells or tissue types. In Australia, an epithelial growth factor is available for "biological wool clipping" of sheep.

Manufacture. EPO is produced in 5,000 L bioreactors (→100) or larger, using recombinant mammalian cells. CHO cells are the preferred cell line. The fermentation process takes up to 30 d. EPO is secreted into the nutrient medium and purified from the culture broth by a series of chromatographic steps (→106). Because an authentic glycosylation pattern (→262) is of key importance for function, the pure recombinant product is thoroughly tested by appropriate methods. In many other growth factors (such as GM-CSF and G-CSF), the glycosylation pattern is of less or no importance for pharmacological function. To produce them, recombinant microorganisms such as *E. coli* can be used.

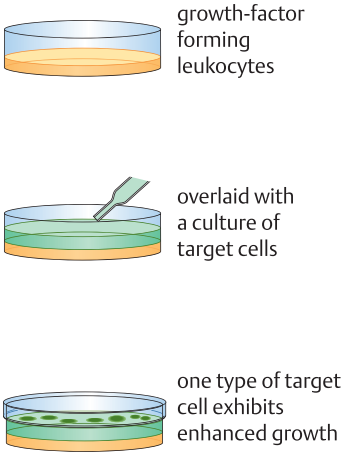
Growth factors in therapeutic use (2013)

product	type of protein	expression system	applications	manufacturer
erythropoietin	glycoprotein, 34 kDa	CHO cells	anemia after dialysis, chemotherapy	Amgen, Roche
granulocyte CSF	glycoprotein, ~ 20 kDa	CHO cells, *E. coli*	immunostimulating agent	Amgen
granulocyte-macrophage CSF	glycoprotein, ~ 18 – 30 kDa	*E. coli*	immunostimulating agent	Immunex
epidermal GF	peptide, 6 kDa	*E. coli*	wool clipping on sheep	BioClip
bone morphogenic protein BMP-7	protein, ca. 100 kDa	CHO cells	bone and cartilage formation	Stryker

CSF = colony-stimulating factor GF = growth factor

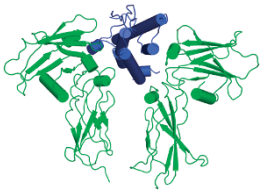
Patent protection for some of these products has already expired, and generic products have entered the market. Bioidenticals are generic products with identical structure, biosimilars are products with similar, but not identical structure (e. g., with a different glycosylation pattern due to manufacture by a different expression host).

Isolation of growth factors



Wool clipping with eGF

EPO

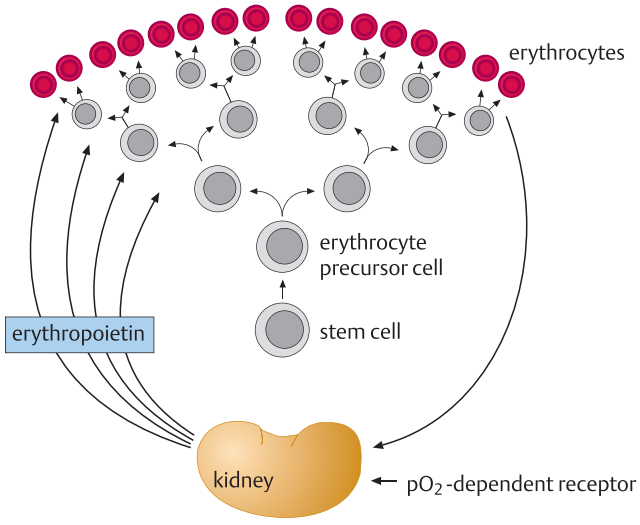


EPO complexed with the extracellular domains of the EPO receptor

green: EPO receptor
blue: EPO

1CN4, resolution 0.28 nm

Function of erythropoietin



EPO and GM-/G-CSF: manufacture and purification

EPO	**recombinant CHO cells**	**cell reactor**	**recovery**
GM-CSF G-CSF	**recombinant *E. coli* cells**	**bioreactor**	**recovery**

quality control

immunoassays, peptide mapping, SDS-PAGE, reverse HPLC, MALDI-TOF, biological activity

Other therapeutic proteins

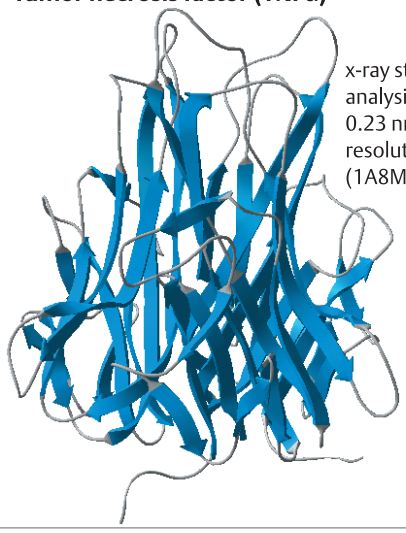
General. Among the hundreds of recombinant proteins that are presently under study for therapeutic applications, the cytokine tumor necrosis factor (TNF), DNase I, and glucocerebrosidase are discussed here.

Tumor necrosis factor. The discovery of TNF relates to early observations that some tumors are halted in their development after the patient has suffered a bacterial infection. It has since been shown that bacterial endotoxins (lipopolysaccharides) stimulate the formation of TNF in activated macrophages, monocytes, natural killer cells (NK cells), and also in liver and brain cells. TNF occurs in two variants, which exhibit similar biological activity although they show < 30 % sequence homology. TNFα (M_R 17.3 kDa, 157 amino acids) is formed by macrophages. Its crystal structure is dominated by an unusually large number of β sheets (→28). Two types of specific TNFα receptors have been found, which occur in quite different cell types. TNFβ (171 amino acids, a glycoprotein) is formed by lymphocytes ("lymphotoxin"), but binds to the same receptors. Its function is less well understood. Originally, the interest in TNFα was focused on the observation that it has a cytotoxic effect on isolated transformed cells which can be further stimulated by interferons. Clinical studies, however, did not corroborate these findings and, to the contrary, showed highly toxic side effects. Indeed, many undesired side effects of the cytokines (inflammation, arthritis, elevated blood pressure, etc.) seem to originate from the formation of TNFα. TNF has also been involved in reactions such as septic shock (which is also initiated by lipopolysaccharides), in cachectic conditions (emaciation) following chronic infections, or in tumor development. It regulates the development of other cytokines (→80), is involved in the development of autoimmune diseases such as rheumatoid arthritis, and plays a role in transplant rejection. This interesting ambivalent role of TNF and its easy production by recombinant hosts (e. g., *E. coli*) have led to significant interest in inhibitors of TNF. This includes monoclonal antibodies (→242) such as infliximab (Remicade®) and adalimumab (Humira®), but also receptor fusion proteins such as etanercept (Enbrel®).

DNase I (Pulmozyme®). Cystic fibrosis (CF) or mucoviscidosis is a genetic disease caused by a frameshift mutation in the *CFTR* gene (cystic fibrosis transmembrane conductance regulator). It affects ca. 30,000 children and adults in the US alone. One in 28 Caucasians is an unknowing, symptom-less carrier of the defective gene. CF causes the body to produce an abnormally thick, sticky mucus, due to the faulty transport of sodium and chloride from within cells lining organs, such as the lungs and pancreas, to their outer surfaces. The thick CF mucus also obstructs the pancreas, preventing enzymes from reaching the intestines. Excessive formation of mucus results in dramatic hindrance of respiration. The viscosity of the sputum is further increased by extracellular DNA originating from leukocytes. A therapy used in cystic fibrosis is the inhalation of an aerosol containing recombinant human DNase I (260 amino acids). The recombinant enzyme is produced in CHO host cells (→98), using a DHFR-containing insert and methotrexate-containing medium for high yields. Since DNase I is inhibited by the G-actin occurring in sputum, mutant enzymes have been engineered (→198) which are not inhibited and thus have 10-fold higher activity in sputum. Cystic fibrosis is a monogenetic disease and thus may lend itself eventually to gene therapy (→304).

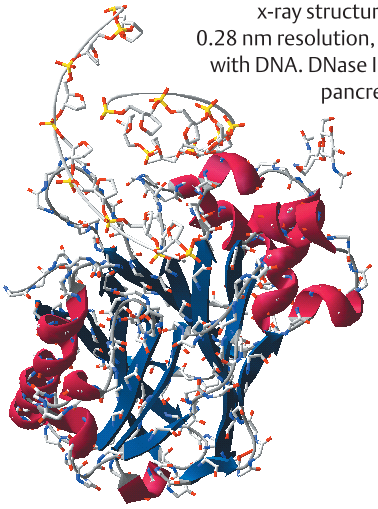
Glucocerebrosidase. Gaucher's disease is an inheritable storage disease. It results from a lack of formation of the enzyme glucocerebrosidase, resulting in the deposition of large amounts of cerebrosides in some cell types. Inheritance is autosomal recessive and coded in one among 40,000 humans, of which 10,000 show symptoms. Based on clinical symptoms, three types of Gaucher's disease are distinguished. Patients of the most frequent variety, Gaucher form 1, have pain in the bones and the digestive tract, but no nerve symptoms. They can be successfully treated by intravenous application of human ß-glucocerebrosidase (Cerezyme™). The enzyme can be obtained from human placenta, but recently, production of the recombinant enzyme in CHO cells (→98) has become preferred. The use of recombinant plant cells for production has also been described. The market volume is ~200 Mill. US$ (2012).

Tumor necrosis factor (TNFα)



x-ray structure analysis, 0.23 nm resolution (1A8M)

DNase I



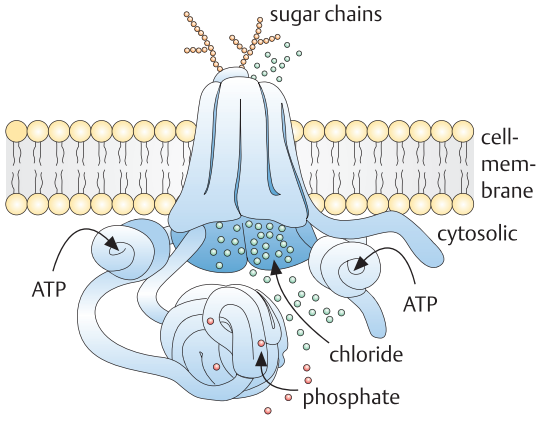
x-ray structure analysis, 0.28 nm resolution, cocrystals with DNA. DNase I from beef pancreas (1ATN)

Cystic fibrosis

reduced secretion of chloride through epithelial cells, mucus formation in the lung.
Caucasians: 1 homozygous in 2000 live births, 1 in 28 heterozygous. Life expectancy for homozygous condition: < 30 y

therapy
recombinant human DNase I as an aerosol

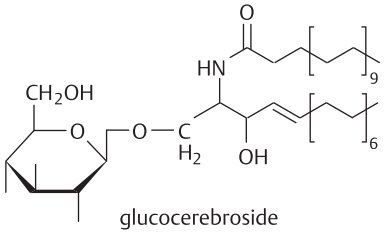
cystic fibrosis membrane regulator (CFTR), a membrane protein made up of a single polypeptide chain of ~2 170 amino acids, coded on chromosome 7 (7q31), is mutated, most frequently (~ 70 %) by a deletion of Phe^{508}



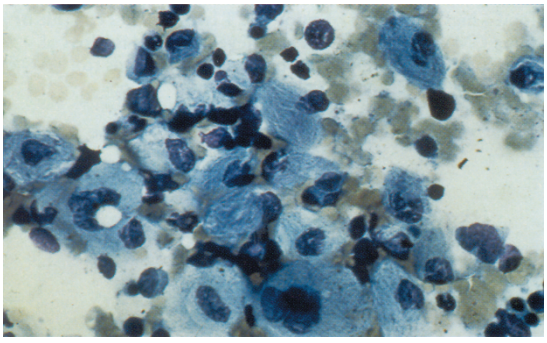
Morbus Gaucher

missing functional enzyme β-glucocerebrosidase (coded on chromosome 1 (1q31), leads to blood anomalies, lung malfunction, bone and nerve defects.

therapy
recombinant human β-glucocerebrosidase



glucocerebroside



bone marrow spread with dyed macrophages (blue), whose form is typical for this microsomal storage disease

CHO cells	cell reactor	recovery	enzymatic transformation
with plasmid pGB20, coding for β-glucocerebrosidase (1 191 bp)	up to 2 000 L	ultrafiltration, chromatography	partial degradation of sugar chains

Monoclonal and catalytic antibodies

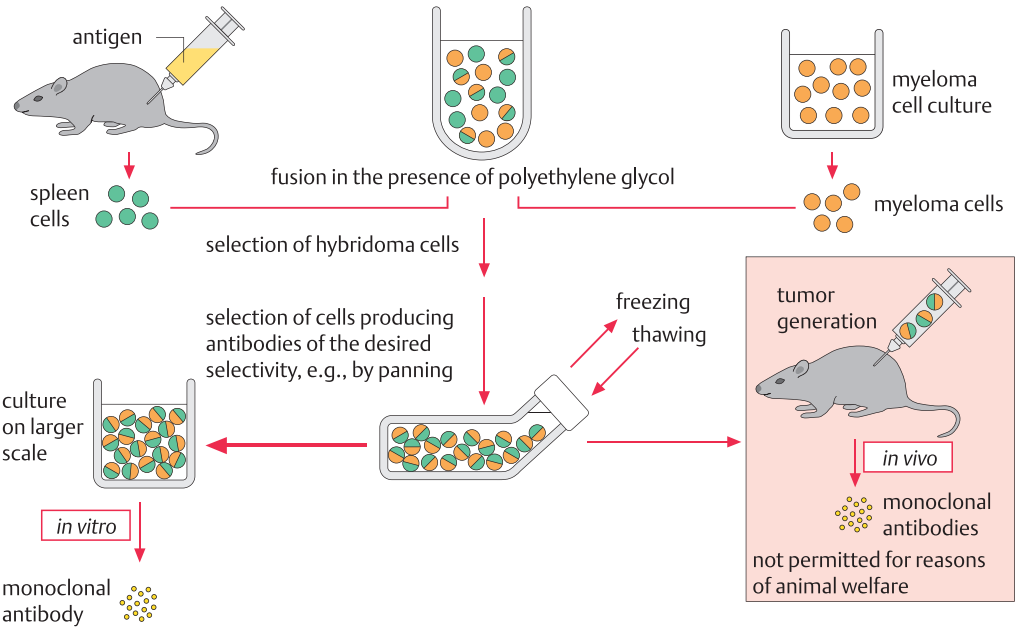
General. In contrast to polyclonal antibodies (→82), which are obtained from immunized animals and consist of a mixture of antibodies directed against the same antigen, monoclonal antibodies are homogeneous: they consist of a single type of antibody with defined selectivity and activity. They are manufactured using hybridoma cells. Monoclonal antibodies have become important biopharmaceuticals and analytical reagents for immunoassays, but polyclonal antibodies are still widely used for medical applications. For example, Privigen® and Gamunex® are pooled fractions of polyvalent IgG antibodies obtained from large groups of blood donors which are used for the therapy of immune defects. Sales for these products were close to 1 bill US$ each in 2011. Also, polyclonal antibodies obtained from immunized animals such as cows or horses are still used in emergency situations such as snakebites.

Hybridoma technology. Antigen is injected into an experimental animal (usually mice, sometimes rats). The spleen lymphocytes are isolated and fused *in vitro* with murine lymphocyte tumor cells (myeloma cells), which can be held in culture and divide indefinitely. Some of the resulting hybridoma cells express antibodies against the antigen on their surfaces and thus can be isolated, using immunoassays and immunological cell cloning procedures. The most productive clones are deep frozen, rendering them stable for many years. This methods allows reproducible production of pure monoclonal antibodies against a nearly unlimited choice of antigens and haptens.

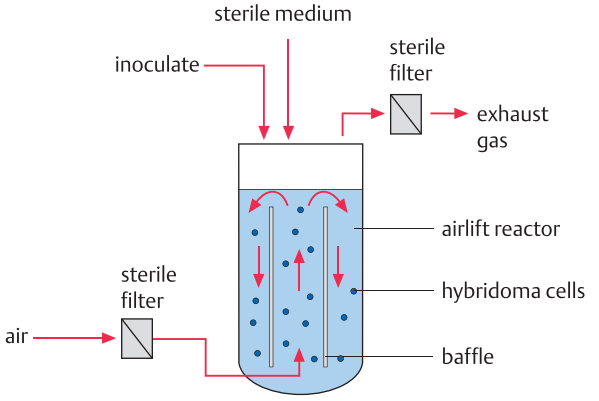
Manufacture of monoclonal antibodies. Hybridoma cells divide and can be propagated in cell culture permanently. A preferred medium is the HAT medium made up from hypoxanthin, thymidine and aminopterin, an antimetabolite for the biosynthesis of nucleic acids. Hybridoma and B-cells have a "salvage pathway" which evades the blockage by aminopterin through de-novo synthesis of purine bases if hypoxanthine and thymidine are present. Monoclonal antibodies synthesized by hybridoma cells are secreted into the culture medium at levels of 10–30 mg L^{-1}. Larger quantities (800– > 1000 mg L^{-1}) are obtained through cultivation of cells in a bioreactor, using complex media. In addition to D-glucose and L-glutamine, fetal calf serum was originally used in the medium, since it contains vital cytokines and growth factors such as lactoferrin. More recently, complex synthetic media have been developed which mimic bovine serum factors (BSF) and contain glucose, glutamine, vitamins, salts, trace elements and recombinant growth factors such as insulin and lactoferrin. Hybridoma cells can be grown under aerobic conditions on solid surfaces, but also have been adapted to grow in suspension. They require the presence of oxygen and CO_2. On a laboratory scale, slowly rotating spinner cultures or roller flasks are mostly used. In industry, bioreactors are applied. They must be optimized for aeration and mixing, since mammalian cells are sensitive to shear forces. Originally, bioreactors with large internal surface areas, such as macroporous beads or hollow-fiber modules were used to protect the cells from mechanical disruption by aeration. Since then, conditions for suspension culture with CHO cells have been developed, using both stirred tanks and airlift reactors (up to 12,500 L) (→100). Fermentations in industry are usually done in fed-batch mode and may yield several grams of monoclonal antibody per liter of culture. In a typical purification procedure, the medium is concentrated by ultrafiltration or diafiltration, followed by binding of the antibody to a protein A column (→106). Further purification may include ion-exchange chromatography and the removal of aggregated antibodies and foreign proteins by gel chromatography.

Catalytic antibodies are obtained from mice immunized with molecules which mimic transition states of an enzyme-catalyzed reaction. Using hybridoma technology and depending on the hapten which has been utilized, monoclonal antibodies are obtained which catalyze reactions that do not necessarily occur in nature, e. g., Diels-Alder reactions. X-ray analysis of the CDR regions of such antibodies show similarities to the active sites of enzymes (→30) that catalyze similar reactions. Thus, the catalytic antibody 17E8, which hydrolyzes formyl-norleucine phenyl esters, forms a catalytic diad of serine-histidine instead of the catalytic triad of serine hydrolases. However, the efficiency of the catalytic antibody (as k_{cat}/K_M) was much lower as compared to the enzyme.

Preparation



Fermentation and recovery



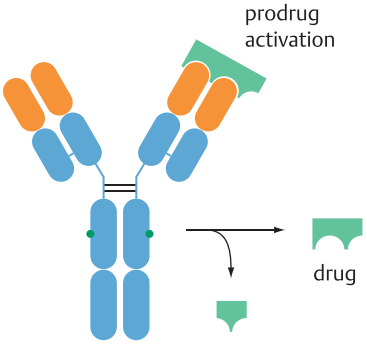
harvested cells
centrifuge, concentrate supernatant

raw protein
protein A or G
anion exchange-, gel chromatography

monoclonal antibodies
confectioning, quality control

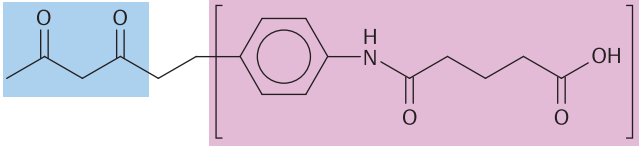
up to > 1 g /L

Catalytic antibodies

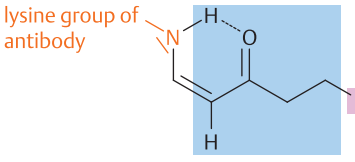


Catalytic aldolase antibody 38C2

hapten: synthetic enamine precursor



hypothetic activated enaminone transition state



(absorbs at 316 nm)

R is coupled to serum albumin

Antibody 38C2 catalyzed > 100 different aldol additions or condensations

Recombinant antibodies

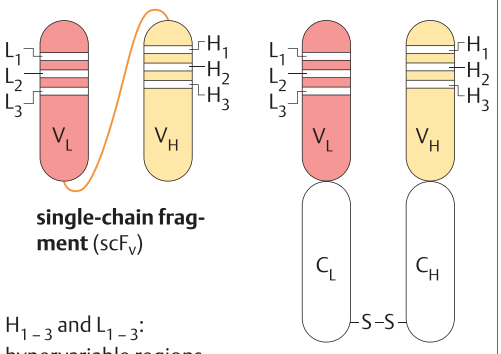
General. Genetic engineering methods can be used to express native or modified antibodies in host organisms. If microorganisms are used, antibody fragments are usually formed, in particular, single-chain (scF_v) and F_{ab} fragments (→82). Complete recombinant antibodies have been obtained from eukaryotic cells (→98), the baculovirus system (→6), transgenic plants ("plantibodies") (→284), and the milk of transgenic animals (→272). Recombinant antibodies have great potential in both diagnostics and therapy. Bispecific and bifunctional antibodies have been studied for targeting drugs bound as antigens to the desired type of cell, e. g., for detoxification, immunosuppression, and cancer therapy. In proteome analysis (→314), specific antibodies have been used to identify and quantify specific proteins after they were separated by 2D electrophoresis, or antibody arrays have been used directly (→316). Catalytic antibodies are being investigated for use in biocatalysis.

Manufacture. The genetic starting material is cDNA, derived either from the mRNA of myeloma cells originating from an immunized laboratory animal or from the mRNA of "naive" B-lymphocytes. After cloning into the host organism, recombinant antibodies or antibody fragments are produced. For example, correctly folded scF_v or F_{ab} fragments can be expressed in the periplasmic space of *Escherichia coli* after cDNA, coding either for a fusion protein or for the separate V_L and V_H chains, has been integrated into a λ vector and competent cells have been transfected with this vector. However, these fragments lack two functional domains: the F_C fragment, which attaches to a receptor, and the glycosylated C_H2 domain. Therapeutic antibodies are thus produced in animal cell culture, e. g., with CHO cells (→98). The production of whole antibodies in transgenic plants is also an active area of research ("plantibodies").

Combinatorial modification. A great advantage of the production of antibodies by recombinant techniques, compared to the hybridoma technique, resides in the potential to prepare very large antibody libraries. These are usually created by a technique called "phage display." (→198) In this method, the whole antibody repertoire of a B lymphocyte is isolated as cDNA, fused with the gene for a viral coat protein, and packaged into M13 expression vectors. After one infectious passage through an *E. coli* host, up to 10^{10} helper phages can be formed which, depending on the cDNA construct used, express on their surface an scF_v or F_{ab} fragment that is coded in their genome. Antibodies of high affinity can be easily isolated from this library by affinity chromatography (→106), and the encoding gene remains within the isolated phage particle. Repeated cycles of mutation and selection based on chain shuffling, error-prone PCR mutagenesis, or mutant strains of *E. coli* have led in a short time to antibodies of remarkable affinity. Thus, by stepwise mutagenesis of the CDR regions (→82) of an antibody fragment (gp120) neutralizing the HIV-1 virus, its affinity could be increased 420 fold to 15 pM. The concept has already been successfully applied to the preparation of human antibodies of high affinity, starting out from a naive cDNA library from human B lymphocytes (→82). Yeast surface display of protein libraries has also been developed. Compared to phage display, the libraries are smaller and hampered by yeast-specific glycosylation.

Antibody arrays. Recombinant antibody libraries are valuable tools for the high-throughput analysis of proteomes (→314, 316). This method is helpful, for example, to study protein patterns of different tissues, or to establish finger-prints of proteins from healthy and diseased tissues, generating biomarkers permitting early diagnoses of diseases. For the construction of such libraries, F_{ab} fragments are mostly used which have usually been prepared using the methods of phage display (→198). Such F_{ab}-fragments (> 10^{10} variants) usually carry an additional and unique binding sequence ("*zip-code*") for the appropriately prepared surface of a glass or polymer chip on which they are bound with high selectivity and at high density (~10^6 variants per mm^2). Surface treatment of the chip prior to binding excludes any nonspecific binding events, leading to an excellent signal to noise ratio upon binding of proteins. Detection of binding events with the proteins is usually done after labeling the complete proteome sample with a fluorescence marker (→84). This procedure allows for a detection limit in the pico- or even femtomolar range. Alternatively, methods such as mass spectrometry which do not rely on labeling are being investigated.

Recombinant antibody fragments



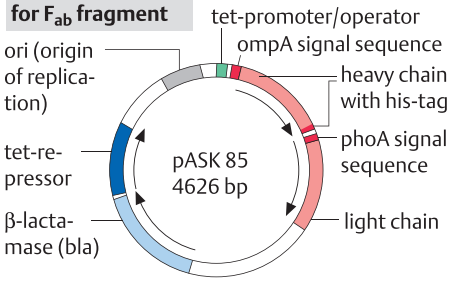
single-chain fragment (scF_v)

H_{1-3} and L_{1-3}: hypervariable regions (CDRs) of the heavy and light chains

F_{ab} fragment (F_{ab})

Expression vector for *E. coli*

for F_{ab} fragment



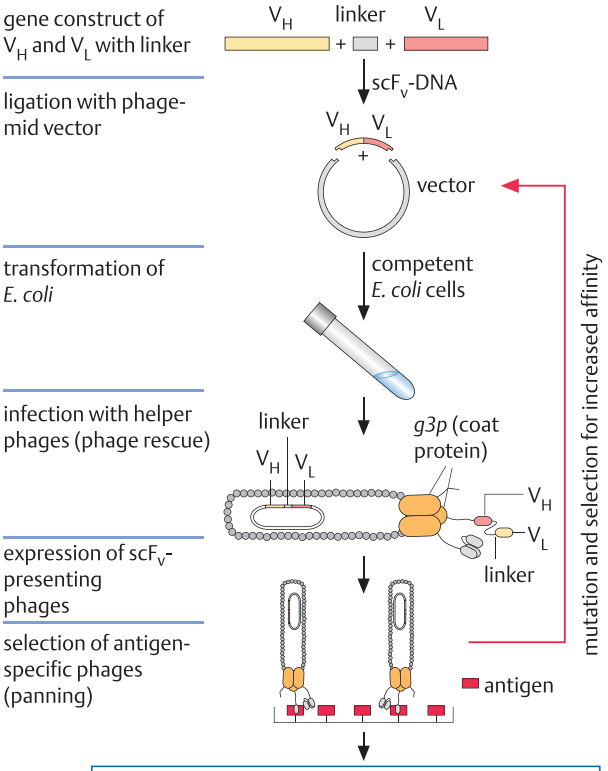
bicistronic expression cassette (1 promoter for the expression of 2 antibody chains)

ompA = outer membrane protein A
phoA = leader-sequence for alkaline phosphatase
bla = Ampicillin resistance

Antibody diversity

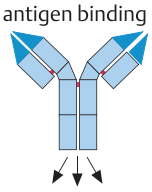
system	repertoire	clonal selection	increase of affinity
B-lymphocytes (immune system)	> 10^8 genes	stimulation of specific B-cells through IgM on their surface	lymphocytes: somatic hypermutation
E. coli	cloned repertoire of lymphocytes plus synthetic genes with random CDR > 10^8 genes	expression of antibody fragments on the surface of phages followed by affinity screening	gene shuffling, mutator strains, error-prone PCR

Clonal selection: phage-display technique



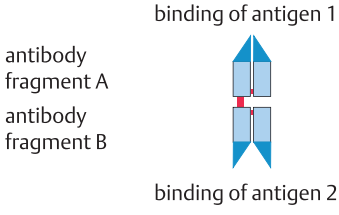
Antibody hybrids

natural IgG



activation of the immune system

bispecific antibody



bifunctional antibody

* enzyme, toxin, radioactive label, biotin binding domain, etc.

Therapeutic antibodies

General. Therapeutic antibodies are among the most successful recombinant protein drugs under clinical application. They are mainly used for cancer therapy and for the treatment of inflammatory diseases. They are parenterally applied and thus can lead to immune reactions. For this reason, chimeric, humanized or human antibodies are preferred over mouse monoclonal antibodies, as they are more similar to human antibodies.

Preparation. For the preparation of chimeric, humanized or human antibodies, three different procedures are being used. Chimeric antibodies are obtained from transgenic mice whose lymphocytes have been modified to generate a human F_c- and mouse ("murine") F_{ab}-fragments (→82). For humanized antibodies, the murine part is limited to insertion of murine CDRs ("complementarity-determining regions") (→82) into otherwise human antibodies. Human antibodies are obtained from the genetic repertoire of human lymphocytes, which are subjected to combinatorial shuffling and selection *in vitro.* The hits are then expanded in cell cultures.

Chimeric antibodies are obtained by splicing genes which code for human F_c-segments with genes coding for the murine F_{ab}-segment. Such products are obtained by combining the human and murine gene segments coding for the desired antibodies in yeast artificial chromosomes (YACs) (→14) and using these YACs to transfect embryonal stem cells (→78) of mice whose antibody-generating systems are genetically disabled. The human part of these antibodies is about 60 %, thus reducing, but not eliminating, immune reactions upon prolonged therapy.

Humanized antibodies are obtained by inserting murine CDRs (→82) into a human antibody-generating system. The synthesis of such antibodies is done in transgenic mice whose capacity to generate heavy and light antibody chains has been genetically silenced and substituted with the repertoire to produce human antibodies. Using this method, murine CDRs are incorporated into the light chains of the human antibodies. The human part of these antibodies is around 90 %.

Human antibodies are formed from the genetic repertoire of naive human B-lymphocytes through clonal selection. This is either done by phage display, a method based on *E. coli*-specific phages, or with baker's yeast using a similar surface display technology, where the gene coding for the yeast surface protein Aga2p is fused with an antibody library. In both cases, highly diverse libraries of human antibodies are prepared, e. g. the company Morphosys' library HuCal contains several billion complete human antibodies. These can be tested for binding events with antigens. An even newer technology (2013) (Ylanthia®) combines the high diversity of humanized mouse antibodies with expression in human B-lymphocytes leading to a library of >100 billion human antibodies of high diversity, good solubility and therapeutic compatibility.

Applications. Among top-selling drugs ("blockbuster drugs") is an impressive number of antibodies or antibody fragments. Adalimumab (Humira®) is similar to its competitor Infliximab (Remicade®), an antibody directed against tumor necrosis factor TNFα (→240). They are being used to treat rheumatoid arthritis, psoriasis, Crohn's disease and ulcerative Colitis (added sales 2012: ca. 16 billion US$). Bevazicumab (Avastin®) suppresses the formation of new blood vessels in tissue (angiogenesis) and is used to treat progressed tumors (sales 2013: > 7 billion US$). Trastuzumab (Herceptin®) binds to the HER2/neu receptor suppressing the growth of those tumor cells which express this receptor. It is used to treat mammary and stomach carcinomas in which these cell types have been confirmed (sales 2013: 7 billion US$). Ranibizumab (Lucentis®), a Fab-antibody fragment, binds to the endothelial growth factor and is used to treat age-related macular degeneration (sales 2013: 4.3 billion US$). Palivizumab (Sinagis®), a monoclonal antibody, is used for passive immunization of children against viral infections.

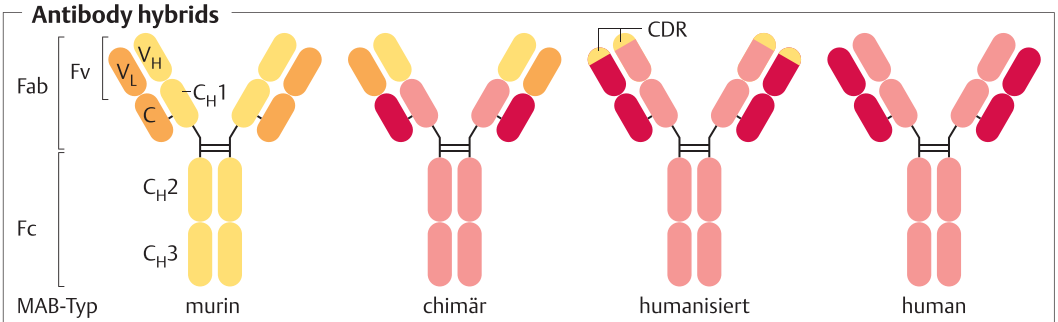
Nomenclature. In view of the large numbers of monoclonal antibodies created for therapeutic or other use, a generic naming system was established by the WHO and adopted by the USA (USAN). It allows the identification of both the source and the target of an antibody. Thus, Tras**tuzu**mab is a humanized antibody (-zu-) directed towards miscellaneous tumors (-tu-), whereas Inf**lixi**mab is a chimeric antibody (-xi-) active in the immune system (-li-).

Diagnostic and therapeutic antibodies (selection)

Generic name*	Brand name	Application	Antibody type	Target protein/function
many		diagnostics (*in-vitro*)	murine	gonadotropin, growth hormon, prostatespecific antigen, herpes, Legionella etc.
	CEA-Scan	tumor diagnostic	murine	carcinoembryonal antigen CEA
Infliximab	Remicade®	rheumatoid arthritis, Morbus Crohn, Psoriasis	chimeric	blocks TNF-α
	Herceptin®	mamma carcinoma, stomach carcinoma	humanized	HER2/new receptor
Alemtuzumab	MabCampath®	lymphatische Leukämie	humanized	lymphocyte CD52-antigen
Bevacizumab	Avastin®	intestinal carcinoma, mamma carcinoma, lung cancer (non-small cell)	humanized	VEGF (vascular endothelial growth factor)
Cetuximab	Erbitux®	intestinal carcinoma, head and neck tumors	chimär	EGF-Rezeptor (EGF = epidermaler Wachstumsfaktor)
Ibritumomab-Tiuxetanb	Zevalin®	non-Hodgkins lymphoma, radioimmunotherapy	murine, doped with Y^{90}	B-lymphocyte CD20-antigen
Oregovomab	OvaRex®	Ovarial carcinoma	murine	CA-125
Panitumumab	Vactibix®	metastasizing colorectal carzinoma	human	EGF receptor
Alemtuzumab	Lemtrada®	multiple sclerosis	humanized	lymphocyte CD52-antigen
Belimumab	Benlysta®	Lupus erythematodes	human	BLys, a cytokine of the TNF superfamily
Omalizumab	Xolair®	heavy asthma	humanized	F_c-part of IgE
	CaroRX®	mouth wash	plantibody	specific binding to *Streptococcus mutans*

murine: from mouse; plantibody: produced in transgenic plant
*nomenclature: murine -omab, primate -imab, chimeric -ximab, humanized -zumab, human -umab

Antibody hybrids



C = constant region, V = variable region, L = light chain, H = heavy chain, CDR = complementarity determining region, yellow/orange: mouse, pink/red = human, Fab = antigen-binding fragment, Fc = constant fragment, Fv = variable fragment

Vaccines

General. Temporary protection against viruses, bacteria, or toxins (e. g., after a snake bite) can be achieved by injection of antibodies specific to the antigen in question (passive immunization) (→80). Much better, and often life-long, protection can be obtained, however, if the immune system is stimulated to produce suitable antibodies by the injection of vaccines (active immunization). After vaccination, the immune system (→80) produces B lymphocytes, which secrete pathogen-specific antibodies; T lymphocytes, which destroy foreign antigenic material; and long-lived B- and T memory cells, which react promptly with an immune response if the antigen should show up again. Vaccines can be whole cells, cell components (e. g., cell wall lipopolysaccharides), or toxic proteins (toxins). Thus, inactivated antigenic materials or weakened (attenuated) viruses (→6) or microorganisms that are still immunogenic but have lost their pathogenic properties are used as vaccines. Attenuated viruses are obtained through a sequence of cell passages. For many decades, a wide number of vaccines have been produced for human protection, e. g., against measles, diphtheria, tetanus, whooping cough, tuberculosis, cholera, and polio. In some countries, farm animals are vaccinated e. g., against foot-and-mouth disease. In spite of this progress, a significant number of diseases have no vaccines available. Thus, a wide range of tropical diseases and AIDS have resisted vaccination. In addition, several infectious diseases which were believed extinct have reappeared, e. g., tuberculosis. The increasing resistance of some diseases to any treatment with antibiotics (→204) adds to this dangerous development. Fortunately, genetic engineering methods have opened a new way to prepare novel and highly pure vaccines.

Vaccine preparation. The conventional method consists of the formulation of inactivated or attenuated antigenic material for subcutaneous, intramuscular, or oral administration. Usually, strains of a pathogen that have lost their pathogenic properties but still stimulate an immune response are used. Alternatively, the pathogen is first cultured in the laboratory and then inactivated by heat or formaldehyde treatment, while keeping its immunogenic properties. For the preparation of vaccines against microbial pathogens or their toxins, the microorganisms are cultured in bioreactors. Until ca. 1970, viruses were preferentially produced in embryonated chicken eggs, and the virus coat protein, after purification from the egg albumen, was used as a vaccine. Today, another preferred method is to propagate the virus in animal cells using animal cell culture technology (→98). Attenuated viruses (→6) are usually used in both processes, and after isolation from hen's eggs or cell culture, they are further weakened or inactivated by treatment with heat or formaldehyde. In view of the potential risks of these fermentation and recovery processes, all steps, including formulation, are carried out under high safety standards (→332). Activity and stability of the product are usually tested in animals ("release tests").

Examples. Tetanus occurs due to infection of a wound by *Clostridium tetani*. During anaerobic growth, this pathogen secretes a neurotoxic protein that is transported by blood to the nerves, resulting in spastic paralysis. For the preparation of tetanus toxin, a hypertoxinogenic strain (Harvard strain) is cultivated in a bioreactor. After growth is complete, the microorganisms are autolyzed, which releases the toxin. After removal of cell debris by filtration, the toxin is inactivated for 4 weeks in a ca. 0.5 % formaldehyde solution, resulting in a "toxoid." This is purified by diafiltration and salt precipitation and absorbed on aluminum salts, thus increasing its immunogenic properties (adjuvant effect). Immunogenicity and tolerability of the lot are tested in animals. To obtain a measles vaccine, animal or human cells in culture are inoculated with a virus strain of low virulence (Edmonton strain). After lysis of the host cells, the virus is isolated by continuous-flow zonal ultracentrifutation and further purified, leading to a freeze-dried or liquid preparation of high stability in storage. Vaccination of farm or domestic animals is another area of rapid developments. Vaccines have become available which are directed against a wide range of animal infectious diseases such as, e. g., foot-and-mouth disease (FMD) in cattle, viral respiratory diseases in horses, cows or pigs, or bacterial infections causing enteritis during mass production of pigs. Veterinary vaccines are estimated to comprise about 20 % of the global vaccine markets.

Basic techniques

passive immunization

application of antibodies

active immunization

1 protective vaccination
2 oral vaccination using
- inactivated pathogens
- attenuated pathogens
- pathogen-specific antigens
- pathogen DNAs

1 for systemic infections
2 for local infections

Gaps in vaccination

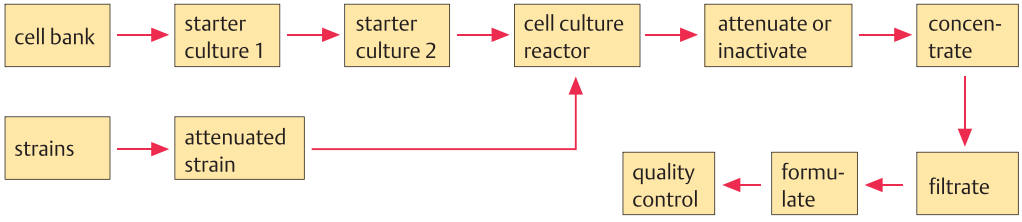
disease	cases per year (millions)	death toll per year (thousands)
diarrhea	> 4 000	> 400
various worm diseases	> 2 000	> 20
respiratory diseases	> 350	> 4 000
malaria	> 300	> 1
schistosomiasis	> 250	> 10
tropical measles	> 44	> 1 000
Chagas disease	> 25	high
tuberculosis	> 6	> 2 000
AIDS	> 5	> 150

Available vaccines (examples)

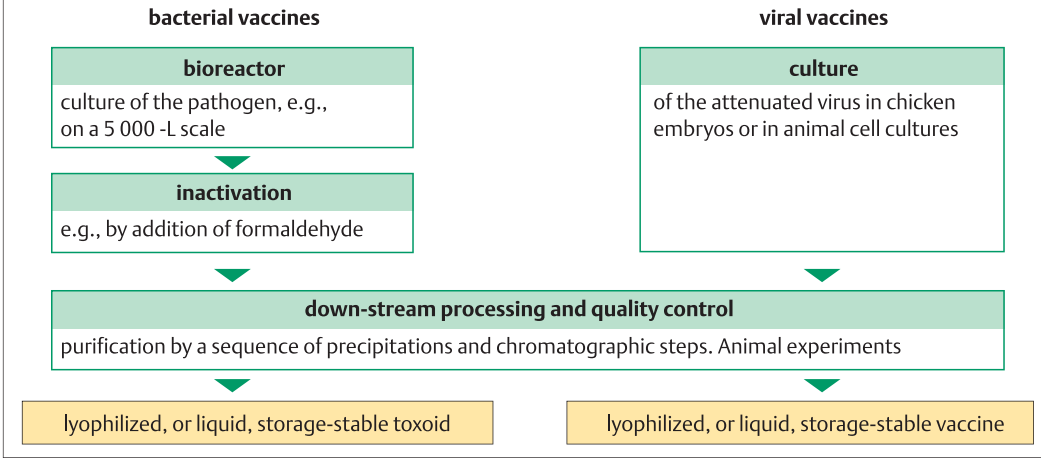
	vaccine	application	manufacture
1	BCG	tuberculosis	attenuated live vaccine from *Mycobacterium bovis*
1	rubella	measles	live vaccine using attenuated *Rubella virus*
2	poliomyelitis	polio	live vaccine using attenuated Poliomyelitis virus
1 2	cholera	cholera	inactivated strains of Vibrio cholerae, live alterated vaccines
2	typhus	typhus	attenuated strains of *Salmonella typhimurium*
1	Haemophilus	meningitis	purified capsular polysaccharide from *Haemophilus influenzae*
1	FMD	foot and mouth disease	aziridine-inactivated FMD virus

*attenuation: weakening through repeated passages in cell culture ("passaging")

Manufacture of a virus vaccine



Fermentation and recovery



Recombinant vaccines

General. Genetic engineering procedures have opened up new possibilities for preparing vaccines. They allow for the manufacture of high purity component vaccines but may also lead to completely new vaccination strategies. Examples are the incorporation of vaccine components into the coat of harmless viruses (→6), the expression of vaccines in transgenic plants (→284) or in the milk of transgenic farm animals (→272) (both procedures would lead to immunization via food intake), and vaccination by direct transfection with DNA or RNA. Up to now, however, only a few recombinant component vaccines have reached the world markets, e. g., vaccines for preventing hepatitis B or diphtheria infections, and a vaccine against Lyme disease.

Strategies. Using genetic engineering, component vaccines can be obtained that are directed towards single-cell components of a pathogen, e. g., against its surface proteins. This procedure implies, however, that the immunogenic components of the pathogen are known. A successful example of this strategy is the production of a recombinant hepatitis B vaccine. In many countries of Asia, hepatitis B is an endemic disease, proceeding undiscovered in 90% of all cases and leading to chronic liver disease in about 5% of the patients. It is estimated that about one billion humans suffer from hepatitis B. To prepare a recombinant vaccine, the surface antigen HbsAg of the hepatitis B virus (→6) was isolated from the blood plasma of infected humans, sequenced, and cloned. It can be expressed in *E. coli*, or, preferentially, in *Saccharomyces cerevisiae* (→14). Purification of the secreted protein is carried out by a sequence of chromatographic steps. A different strategy for recombinant vaccines is the preparation of genetically attenuated host strains. For example, *Vibrio cholerae*, the strain responsible for cholera, normally produces cholera toxin, a fusion protein with adenylate cyclase activity. After secretion within the small intestine, it leads to the formation of cAMP, resulting in a massive loss of liquids and electrolytes, with a clinical picture of violent diarrhea. Using genetic engineering techniques, a deletion mutant of *V. cholerae* was prepared that lacks adenylate cyclase activity but shares its immunogenic properties with the pathogenic strain and thus can be safely used for vaccination. As a third strategy, vector vaccines are being proposed: vaccination with viral DNA that has been modified so as to code for the desired immunogenic proteins but lack all pathogenic elements. As a vector, the cattle pox virus *Vaccinia* was chosen because it is highly infective but completely harmless to humans. Viral antigens were successfully engineered into the *Vaccinia* (→6) genome, resulting, after infection, in immunization against the G protein of the rabies virus, the hepatitis B virus surface antigen, the NP- and HA proteins of influenza virus, and other antigens. The concept, however, is not considered safe enough, in particular for infants or immunosuppressed patients. Local or pandemic outbreaks of highly pathogenic flu in poultry since 1959 have alarmed the general public and the health authorities. Influenza-A virus type H5N1 and variants were shown to be the responsible vectors. Using reverse genetics methods, vaccines against the H5N1 virus were generated based on the viral hemagglutinin and/or neuraminidase antigens; they protected both mice and chicken against infection.

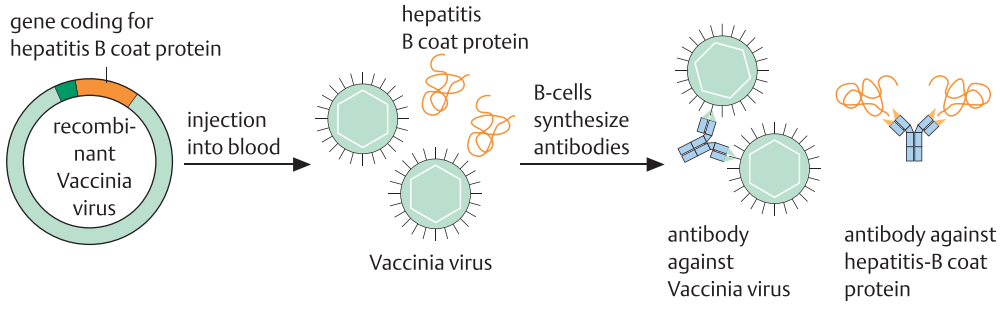
Transgenic plants are being proposed as a vaccination program in developing countries (eatable vaccines, e. g., transgenic bananas). Since the vaccine would be taken up by the gastrointestinal tract (as occurs with oral vaccination), its efficacy depends on its stability during gastrointestinal passage and its transfer through the mucous membrane, where it must stimulate the immune system of the small intestine to produce antibodies. This concept also raises a number of regulatory questions, e. g., of consistent production and the behavior of the vaccine protein during ripening, decay, or processing of fruits or other food products.

DNA vaccines. After the injection of DNA coding for the surface structures of the malaria pathogen *Plasmodium falciparum* into the spleen of mice, the vaccinated animals produced antibodies against this parasite. Similar experiments with the genomic digest of *Mycobyceterium tuberculosis*, the tuberculosis pathogen, also led to a T cell response and allowed for the identification of gene products leading to an immune response. In both studies, the antigen-specific DNA was integrated into plasmids. This interesting new method is still in its infancy.

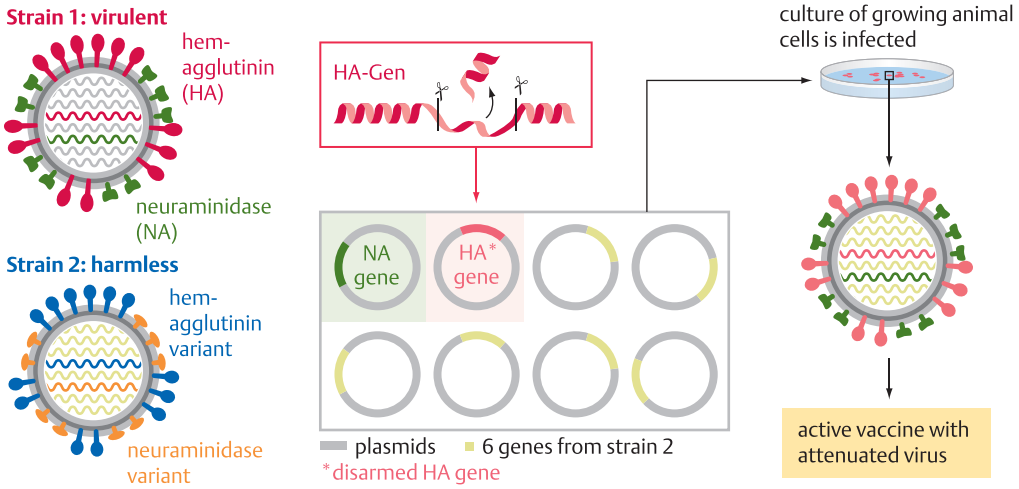
Recombinant vaccines (selection)

		antigen	status
viruses	hepatitis B	surface antigens	registered
	Herpes simplex type 2	surface antigens	clinical studies
	rabies vaccine	surface antigens	not registered
	yellow fever virus	surface antigens	preclinical studies
	AIDS virus	surface antigens	clinical studies
bacteria	*Streptococcus pneumoniae*	polysaccharide conjugate	registered
	Clostridium tetani	tetanus toxin	not registered
	Mycobacterium tuberculosis	surface antigens	clinical studies
parasites	*Plasmodium falciparum*	(malaria)	clinical studies
	Trypanosoma sp.	(sleeping sickness)	clinical studies
	Schistosoma mansoni	(bilharziosis)	clinical studies

Vaccination by recombinant Vaccinia virus



H5N1 avian flu vaccine



Fermentation and recovery of recombinant hepatitis B vaccine

bioreactor	recovery	
recombinant *S. cerevisiae* expresses plasmid-coded rHBAg protein	by precipitation, diafiltration, chromatography	rHBAg vaccine

complex quality control (absence of pathogens, allergens, etc.)

Steroid biotransformations

General. Steroids are a class of chemicals for which biotransformations (→164) have been successfully applied.

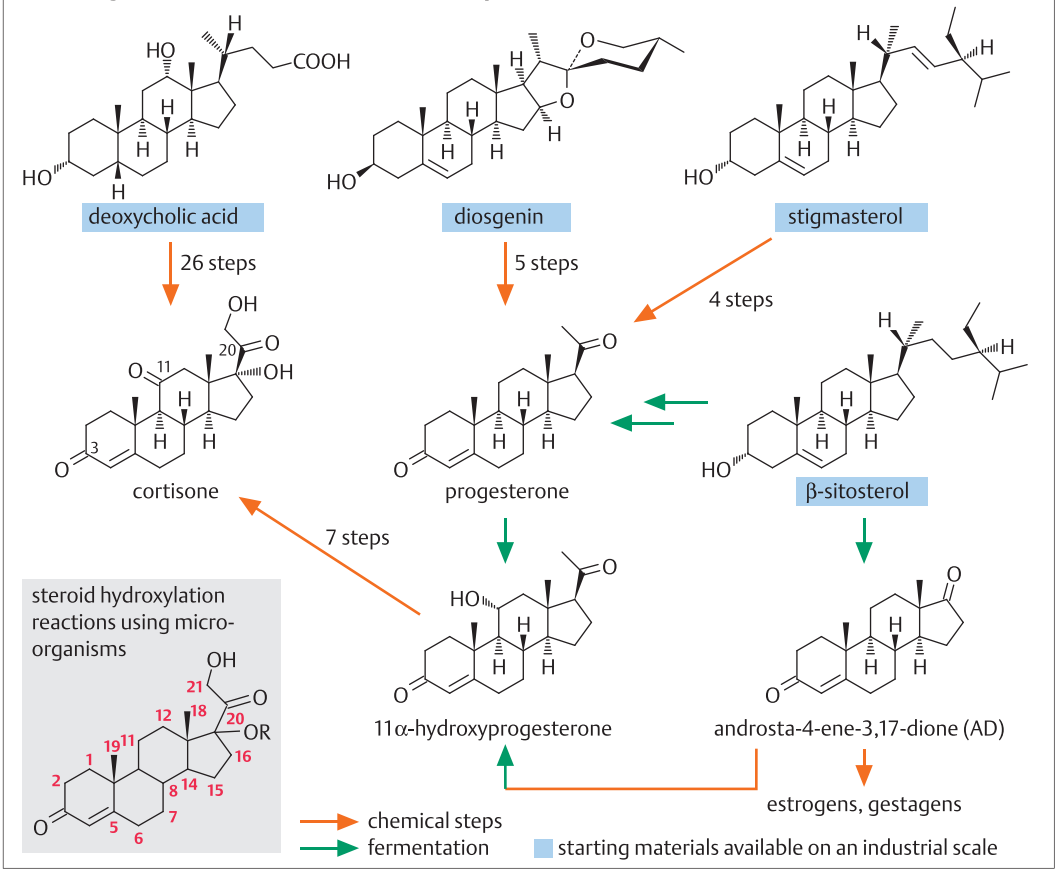
Steroids. This large structural group comprises over 10,000 natural and synthetic compounds. Many of them are used in pharmacology. Examples are vitamin D (calciferol) (→134), anti-inflammatory drugs (corticosteroids), ovulation inhibitors (estrogens and progestins), antiarhythmics (digitalis glycosides), and diuretics (spironolactone). Several types of biotransformation reactions are used in their industrial manufacture. Important examples are the sidechain degradation of β-sitosterol to androsta-4-ene-3,17-dione (AD), androsta-1,4-diene-3,17--dione (ADD), and the 11β-hydroxylation of cortexolone ("Reichstein S"). Desaturation of AD to ADD is also used industrially. The total synthesis of pregnenolone from sugar has been achieved using a recombinant yeast carrying several foreign genes. Ursodeoxycholic acid is used to dissolve cholesterol gallstones and to treat primary biliary cirrhosis; the chemical synthesis from cholic acid can be shortened by regio- and stereospecific enzymatic isomerization of hydroxy-groups in cholic acid, using appropriate dehydrogenases and a cofactor regeneration system.

Sidechain degradation. For a long time, diosgenin, a natural compound isolated from a Mexican root, was a key substance in the industrial synthesis of steroids. Bile acids from animal gallbladders and stigmasterol, a side product of vitamin E production from soy oil, were other raw materials. Because chemical synthesis of corticosteroids, estrogens or spironolactone from these starting materials requires a large number of steps, sidechain degradation of plant sterols (e. g. β-sitosterol, isolated from rape or soy seeds) eventually became an attractive industrial alternative. Actinomycetales such as *Mycobacterium*, *Nocardia*, *Arthrobacter*, and *Corynebacterium* have this capacity. They can degrade the sidechain of phytosterols directly to steroid intermediates such as androsta-4-ene-3,17-dione (AD) or androsta-1,4--diene-3,17-dione (ADD), which are starting materials for the synthesis of estrogens and progestins and can also be used to add sidechains chemically, at position 17, which lead to the corticosteroids. AD can be desaturated to ADD with resting cells of *Arthrobacter simplex* or several Mycobacteria.

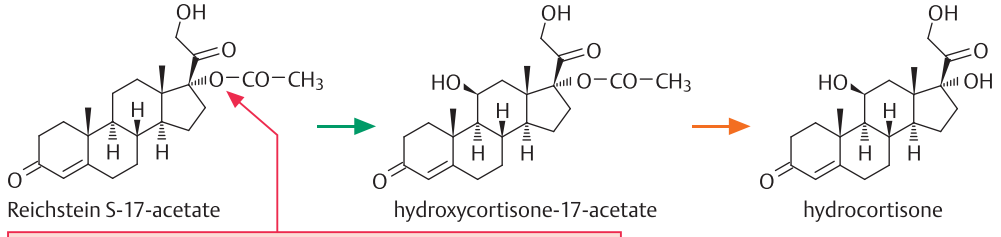
Hydroxylations. Based on many decades of microbial screening, today's culture collections contain microorganisms that can hydroxylate more or less selectively at nearly all positions of the steroid ring. Among the most important industrial hydroxylation steps are those in position 9α of AD and 11α of progesteron, leading to simplified routes for the synthesis of glucocorticoids. 11β-hydroxylation by *Curvularia lunata*, a mold, leads directly to precursors for prednisolone-type corticosteroids. To prevent side reactions of the hydroxylating enzyme, cytochrome CYP11B (it would hydroxylate at positions 7α and 14α in addition to 11ß), the 17α-acetate ester of Reichstein S is used as a substrate. In this system, the desired biotransformation step proceeds with high regio- and stereoselectivity, because undesired side-reactions are sterically hindered. As water solubility of both substrate and product are low, this transformation reaction proceeds on the suspended substrate *via* its dissolved fraction over several days.

Biosynthesis of pregnenolone from sugar. A recent approach consists of constructing a recombinant yeast strain that can form pregnenolone from sugar. During growth, *Saccharomyces cerevisiae* forms ergosterol as a component of its membrane. In the recombinant strain, oxidative degradation of this mycosterol was first suppressed by switching off the gene for Δ22-desaturation. Cloning and functional expression of three bovine genes involved in steroid metabolism and of a gene for a δ7-reductase from the plant *Arabidopsis thaliana* on chromosomes XIII, XV, and III of the yeast resulted in a strain that forms pregnenolone from D-galactose. Further cloning of a 3β-hydroxysteroid dehydrogenase into this mutant led to the formation of progesterone. The pathway was completed by the functional expression of bovine 11β, 17α-, and 21-hydroxylases, leading to the formation of hydrocortisone from sugar in one fermentation step. After more than a decade of basic research, yields have recently become economically competitive and from the end of 2014, this elegant process using a synthetic biology approach (→320) will be commercialized by Sanofi-Aventis.

Starting materials and transformation steps

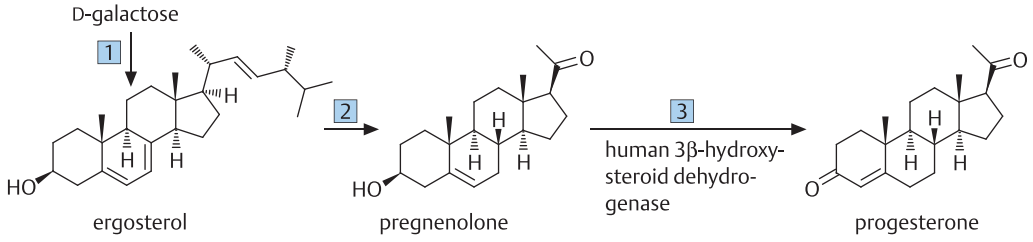


11β-Hydroxylation of Reichstein S using *Curvularia lunata*



undesired hydroxylation at positions 7α and 14α is suppressed by acetylation of the substrate in position 17

Steroid synthesis using recombinant yeast



1 natural metabolism of *Saccharomyces cerevisiae*
2 cloned genes from beef and *Arabidopsis thaliana*
3 cloned human gene

Diagnostic enzymes

General. Enzymes from all six classes (→166) are used for analytical or diagnostic purposes. Their main advantage is their high substrate specificity, which allows selective detection of even a minor component in a complex mixture. To guarantee the highest possible specificity, the enzyme preparation must be free of other enzymes that would interfere with the intended application. Thus, analytical and diagnostic enzymes are of considerable purity (→106). Many analytical determinations with enzymes can be carried out conveniently by use of commercial reagent kits and automated laboratory robots or with test strips. Enzymes can also be used as reporter groups (→84) indicating binding events, e. g., by linkage to antibodies or DNA fragments; if an excess of enzyme substrate is present in such assays, the signal resulting from a binding event is amplified considerably. Major applications of analytical enzymes are in enzyme diagnostics (enzyme tests with automatic laboratory systems, point-of-care test kits, glucose biosensors (→258), enzyme immunoassays etc.) and enzyme tests for food analysis.

Principles of measurement. Usually, photometric, fluorimetric, or luminometric assays are used for measuring enzyme reactions (→84). If direct measurement of the substrate or product of the primary enzyme reaction is not feasible, auxiliary enzyme cascades are often used. Frequently, they include a NAD(P)H-dependent dehydrogenase as indicator enzyme; NAD(P)H formation or consumption can be monitored easily by photometric measurements at or near 334, 340, or 366 nm. These measurements are usually performed with μL amounts of sample (e. g., blood). Thus, in addition to providing extremely pure enzymes, the development of highly precise pipettes, optical equipment, and other devices adapted to such small volumes (μL-technology) was another milestone in establishing enzyme technology.

Methods of determination. (→030 These include three procedures: 1) end-point determinations, 2) kinetic methods, and 3) catalytic methods. In *end-point determinations*, the turnover of a substrate or cofactor in a reaction is measured after its completion. Examples are the determination of ethanol with alcohol dehydrogenase, of lactate with lactate dehydrogenase, or, in a coupled reaction, the decarboxylation of citric acid to oxaloacetic acid by citrate lyase, and reduction of oxaloacetic acid to L-maleic acid by L-malate dehydrogenase, monitoring the concomitant oxidation of NADH to NAD^+. End-point methods require at least several minutes: reaction velocity increases with enzyme concentration and with a low K_m or high v_{max} value. An example of the use of an auxiliary reaction is the determination of D-glucose with hexokinase, followed by the formation of NADPH in the coupled reaction of glucose-6-phosphate dehydrogenase. With the same reaction, but using a kinetic method (→30), glucose levels can be determined much faster. Kinetic methods do not require that a reaction proceed until completion; rather, its initial velocity is measured, which is linearly proportional to the rate-determining enzymatic step, provided the substrate concentration is low (< 1/10 of its K_m value). Reaction conditions and measuring time intervals must be kept strictly constant in this type of assay. Thus, kinetic methods are the method of choice with clinical laboratory robots. *Catalytic methods* enable an even lower limit of detection, since the substance being analyzed is used as the rate-limiting factor in a cyclic multienzyme reaction leading to the continuous consumption of a regenerating cofactor. An example is the determination of coenzyme A with the combined reactions of phosphotransacetylase, citrate synthase, and malate dehydrogenase.

Preparation and properties. Analytical enzymes are produced in small quantities, but at high purity. Often, intracellular enzymes (→166) are applied, which occur in low concentrations and which must be isolated from disrupted cells in high quality and with little or no side activities (→106), using highly diverse chromatographic methods of enzyme purification. Today, recombinant enzymes dominate this area, since they can be produced much more easily in larger quantities and without side-activities. Recombinant enzymes can be further optimized for a given purpose by protein engineering (→198). Apart from specificity and purity, another important issue with such enzymes is their stability during shipment and storage. Typically, activity losses of < 20 % per year at temperatures up to 40 °C can be tolerated. For this purpose, stabilizers such as sugars or glycerol are often added during the finishing process.

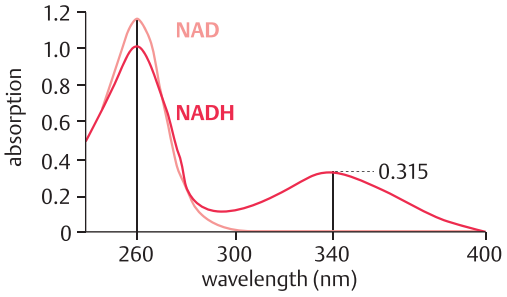
Diagnostic and analytical enzymes

enzyme	EC number	analyte	auxiliary analyte
alcohol dehydrogenase	1.1.1.1	ethanol, other alcohols, aldehydes	NADH
glucose oxidase	1.1.3.4	glucose	dye
pyruvate kinase	2.7.1.40	phosphoenol pyruvate, ADP	
creatine kinase	3.5.3.3	creatine	
citrate lyase	4.1.2.6	citric acid	NAD^+
mannose-6-phosphate isomerase	5.3.1.8	mannose	
succinyl-CoA synthase	6.2.1.4	succinate	

End point determinations

measurement technique	detection limit
photometry	
– end point determinations	1 – 10 µm
– kinetic methods	0.1 – 1 µm
– catalytic methods	1 – 10 nm
fluorimetry	
– end point determinations	1 – 10 nm
– catalytic methods	1 – 10 fm
luminometry	1 – 1 000 pm

determination of NADH



End point determinations

a

ethanol + NAD^+ —(alcohol dehydrogenase)→ acetylaldehyde + NADH + H^+

lactate + NAD^+ —(lactate dehydrogenase)→ pyruvate + NADH + H^+

direct determination of a substrate, a product, or of NADH

b

glucose + ATP —(hexokinase)→ glucose-6-phosphate + ADP

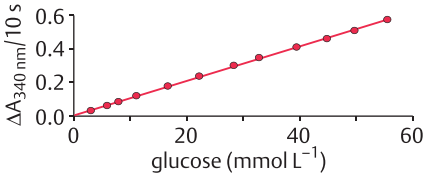
glucose-6-phosphate + $NADP^+$ + H_2O —(glucose-6-phosphate dehydrogenase)→ 6-phosphogluconate + NADPH + H^+

analyte detected species

end point determination with coupled reaction: glucose-6-phosphate dehydrogenase is highly specific, the low specificity of hexokinase remains without consequences

Kinetic assays

kinetic determination of glucose with hexokinase and glucose-6-phosphate dehydrogenase; formation of NADPH after 10 s in an automated system



$v = v_{max} \times [S]/K_m + [S]$ (Michaelis-Menten equation): for substrate concentrations $S >> K_m$ the velocity v of an enzyme reaction is linearly proportional to the substrate concentration

dependence of enzyme dosage on K_m

analyte	enzyme	K_m (µmol L^{-1})	V/K_m (= 1 mL min^{-1})*
ADP	adenylate kinase	1,600	1,600
glucose	hexokinase	100	100
glycerol	glycerol kinase	50	50
uric acid	uricase	17	17
fumarate	fumarase	1.7	1.7

* given for the following condition: the higher the K_m, the higher dosage of enzyme is required, if 99 % of the substrate should be transformed in a few minutes (V/K_m should be on the order of 1 mL min^{-1})

Enzyme tests

General. Since about 1950, the enzymatic determination of metabolites and the measurement of enzyme activities in body fluids have revolutionized medical diagnosis. A little later, immunoassays (→260) and, more recently, DNA assays have become indispensable tools of the physician. In enzyme diagnostics, low-molecular-weight compounds such as lactic acid or glucose are determined in blood or serum by means of appropriate enzymes of high selectivity. Alternatively, suitable indicator reactions are used to determine the concentration of enzymes in blood serum for the diagnosis of destructive processes in individual organs. Enzyme tests have also found entrance in food analysis, in the monitoring of fermentation processes (→96), and in environmental protection. If dehydrogenases are employed, the change in the spectral absorption or fluorescence of $NAD(P)^+$ and NAD(P)H can be used ("optical tests"). Using this principle with suitable dehydrogenases, glucose and ethanol can be measured directly, whereas many other substances, such as fatty acids or glycerol, can be determined quantitatively by coupling two or more different enzymatic reactions. For practical reasons, it is advantageous to use as few auxiliary enzymes as possible. In food analysis, a range of sugars (glucose, galactose, maltose) and acids (citrate, malate) are determined enzymatically. Quasi-continuous determination of glucose is important in most microbial fermentations, and similar monitoring of glucose and lactate is of importance in animal cell culture (→96). The inhibition of an enzyme activity may also be used as an analytical tool: for example, measurement of the inhibition of isolated acetylcholine esterase, a key enzyme in neurotransmission, can be used to indicate the presence of nerve gases or of organophosphate or carbamate pesticides.

Determination of enzyme activities. The measurement of enzyme activities in the serum of a patient is an important tool for the physician. When cells are destroyed during a disease (heart infarction, hepatitis, cirrhosis, pancreatitis, etc.), enzymes are released into the bloodstream. Thus, damage to the cells of various organs (heart, muscle, liver) and their compartments (membranes, cytoplasma, mitochondria) can be monitored by measuring enzyme activity in the blood. An organ-specific distinction (differential diagnosis) is possible, since the cells of different organs often release different enzymes or different enzyme subtypes (isoforms) of an enzyme. Thus, liver diseases can be monitored by measuring transaminases; pancreatic inflammation by α-amylase and pancreatic lipase PL; and the destruction of heart muscle cells by a creatine kinase isoenzyme (the latter method has already been widely substituted by an immunoassay for cardiac troponin T (TNT)) (→260). Their determination follows a kinetic regime based on the use of substrates that show little or no cross reactivity with other enzymes.

Laboratory automation. Clinical enzyme assays were originally performed manually, using individual sample preparations, incubation procedures, and determinations using simple photometers with filters of suitable wavelengths. Now, the tremendous increase in clinical enzyme tests has led to almost complete automation in central diagnostic laboratories, where all pipetting, incubation, and measurement steps are carried out by laboratory robots, and barcodes or other procedures are used for sample identification. In modern robots, measurement principles such as enzyme and immunological tests (ELISA) are integrated.

Test strips. Occasionally, an assay may be too time-consuming for "point-of-care" use (e. g., daily self-tests by diabetics, urgent tests in hospital emergency rooms). In such situations, diagnostic test strips are often used. They are composed of several solid separation and reaction layers ("dry chemistry"), with which the sample is prepared and the indicator reaction proceeds once a solution (e. g., a drop of blood) is added. The enzymes used here are often oxidases; in the presence of the auxiliary enzyme peroxidase, their primary reaction product, H_2O_2, oxidizes a leuko dye forming a colored compound, the amount of which is proportional to that of the component to be measured in the sample. Test strips can be analyzed semiquantitatively by visual comparison with a graded color chart, or quantitatively with a reflectance photometer. In food analysis, test strips have been developed for, e. g., determining the freshness of fish (Japan); the principle used was the liberation of amines, as determined by their oxidation *via* amine oxidase coupled with the formation of H_2O_2.

Frequent enzyme tests (selection)

	typical analyte	relevance
clincial diagnostics	glucose	diabetes
	triglycerides, cholesterol	blood fat, risk of atherosclerosis
	uric acid	gout
	alkaline and acid phosphatase	tumors
	transaminases	liver diseases
	creatine kinase	heart infarction
food analysis	sugars (glucose, maltose)	quality
	acids (citrate, malate)	quality, counterfeits
bioprocess control	glucose, lactate	microbial and cell culture fermentation
environmental monitoring	inhibitors of acetylcholine esterase	presence of organophosphates, carbamates

Differential diagnosis of various organ malfunctions

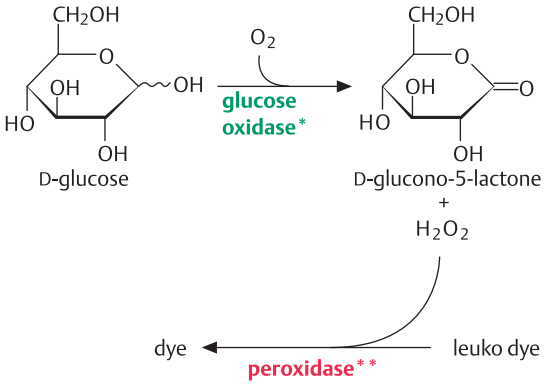
ASAT		~ 50 U/L organ?	
ALAT		~ 20 U/L organ?	~ 100 U/L liver? gall bladder?
CK	~ 15 U/L organ?	~ 250 U/L heart	~ 100 U/L skeletal muscles
LDH	~ 150 U/L organ?	~ 2 500 U/L blood cells	
AP	~ 120 U/L organ?	~ 500 U/L bone?	liver? gall bladder?
γ-GT	~ 75 U/L liver	~ 15 U/L bone	~ 750 U/L liver? gall bladder?

ASAT aspartate aminotransferase
GLDH glutamate dehydrogenase
ALAT alanine aminotransferase
CHE choline esterase
CK creatine kinase
LDH lactate dehydrogenase
AP alkaline phosphatase
γ-GT γ-glutamyl transpeptidase

key enzymes for the diagnosis of organ diseases

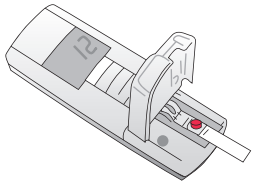
liver, gall bladder: ALAT-GLDH-γ-GT-CHE
heart: CK-LDH (isoenzyme-1)-ASAT
skeletal muscle: CK-GLDH
bone: AP
blood: LDH
pancreas: α-amylase, lipase

Test strip (example: blood sugar)

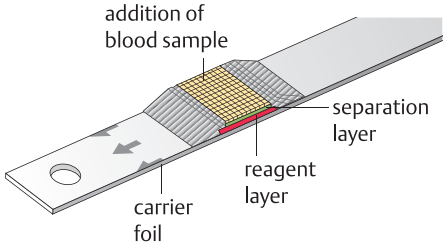


*glucose oxidase from *Aspergillus niger*
**peroxidase from horseradish

hand-held photometer for determination of blood sugar



composition of the test strip



Biosensors

General. In a biosensor, biological recognition by an enzyme, an antibody, aptamer, DNA, a microorganism, etc. is linked to a physical transducer such as an electrode, a fiber-optic device, or a piezo crystal. In spite of intense research and numerous concepts, only a few biosensors have been commercially successful. The market volume for biosensors in 2011 was ~ 9 billion US$, and dominated by various types of glucose biosensors.

Electrochemical biosensors. Oxidases or hydrolases are used for electrochemical biosensors. By hydrolyzing their substrate, hydrolases cause a change in pH which can be monitored by ion-selective electrodes (pH electrodes) or field-effect transistors. Oxidases generate H_2O_2 and consume O_2. Both substances can be analyzed with an amperometric electrode, such as an oxygen electrode. By using so-called mediators, which transfer electrons to the flavine group of the oxidase, the oxidation potential can be significantly reduced. Thus, although the oxidation of H_2O_2 requires a potential of + 400 mV vs Ag/AgCl, dimethylferrocene is oxidized at + 100 mV, eliminating the risk of a false-positive signal if, e. g., L-ascorbic acid (normal potential $\varepsilon = +170$ mV) is present in the sample. In 3rd generation biosensors, mediators are replaced by organic conducting materials based on charge-transfer complexes which perform direct electron transfers from the electrode to the flavine group. The most commercially successful example of an enzyme electrode is the glucose biosensor, which is often used in hospitals as a point-of-care item for rapid glucose determinations, but is mainly used in pocket format for self-testing by diabetic patients (→222). Glucose sensors can also be used for monitoring glucose consumption in bioreactors (→96). Miniature implantable glucose sensors are available for diabetic patients but must be calibrated daily and replaced within a few days due to their insufficient histocompatibility, rendering them yet unsatisfactory for controlling an automatic insulin pump. If a pure or mixed culture of aerobic microorganisms is immobilized on top of an oxygen electrode, respiration of the culture can be continuously monitored. This concept has been commercialized for the continuous measurement of wastewater BOD (→286), resulting in data acquisition in minutes instead of days. Enzyme immunoassays can be based on electrochemical reporter reactions instead of color reactions (→84). The intercalation of DNA with electrochemically active reagents such as daunomycin allows the electrochemical determination of hybridization events (→302). Neither concept has yet been brought to market.

Optical biosensors. If an antibody interacts with an antigen, its mass increases, which can be observed as a change in the surface properties ("evanescent field") of an optical transducer. Instruments based on this principle have found good acceptance in the research market, since they allow a highly sensitive (nM) kinetic determination of antigen/antibody interactions. Other optical biosensors are based on fluorescence quenching by oxygen of polyaromatic chemicals, such as pyrene, providing a substrate-specific signal in the presence of an oxidase system (e. g., glucose oxidase) which consumes equimolar amounts of oxygen upon substrate oxidation (e. g., glucose). In a similar setup, pH changes can be monitored by the fluorescence of fluorescein which is pH dependent. If, for example, penicillin amidase is coupled to such a sensor, the formation of 5-APA from penicillin G can be monitored. The term "optode" or "optrode" is used in this context.

Flow injection analysis (FIA). Although FIA in a strict sense is not a biosensor technology (the biological components and transducer are usually separated from each other), this method is extremely useful for enzyme, immuno, and DNA assays. It combines analysis with automated liquid handling and thus can be advantageously applied for repeated assays of one or a few analytes. The FIA concept has been successfully transferred to the microsystem and nanotechnology fields.

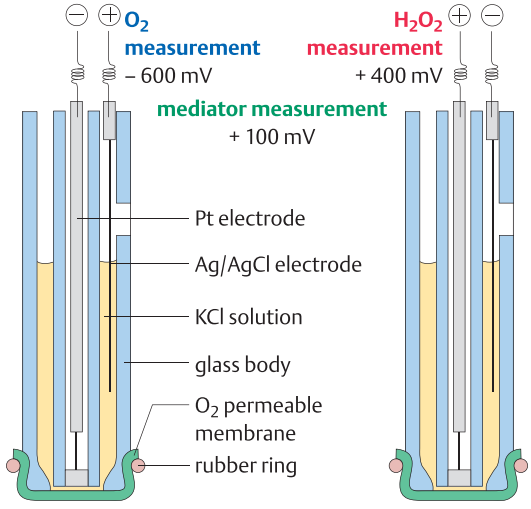
Natural biosensors. The chemoreceptors of bacteria and the sensory organs of higher animals are interesting examples of natural biosensors. These natural biosensors can reliably and quickly analyze highly complex mixtures of chemicals (e. g., rose fragrance, wine aroma). A technical simulation of natural biosensors resides in chemometric analysis based on neural networks. Although this technology is being investigated for use in biosensors, the results so far are a far cry from the performance of natural sensory systems.

Biosensors

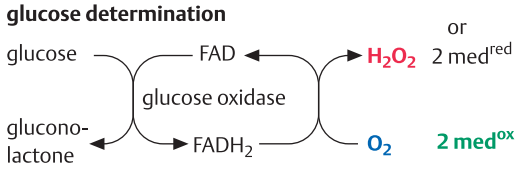
	biological component	transducer
enzyme electrodes		
amperometric	mostly oxidases	O_2 electrode
potentiometric	mostly hydrolases	ion-selective electrode
enzyme FET (field effect transistor)	mostly hydrolases	field effect transistors
microbial sensors	microorganisms	O_2 electrode ion-selective electrode
piezosensors	antibodies	piezoelectric quartz crystal
optical sensors	antibodies, DNA	optical fiber with surface plasmon resonance or grating coupler

sample preparation is often done by liquid handling, e.g., flow injection analysis (FIA); signal processing is usually electronic, complex signals being analyzed by pattern recognition with a neuronal network

Oxygen electrode



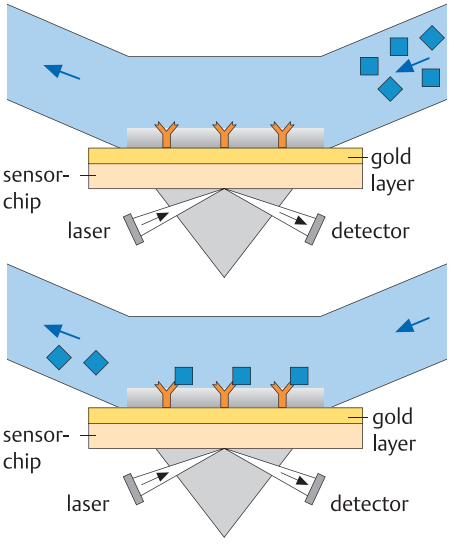
glucose determination



$FAD/FADH_2$: cofactor flavin adenine dinucleotide
med^{red}/med^{ox}: mediator, e.g., dimethylferrocene

Optical biosensor

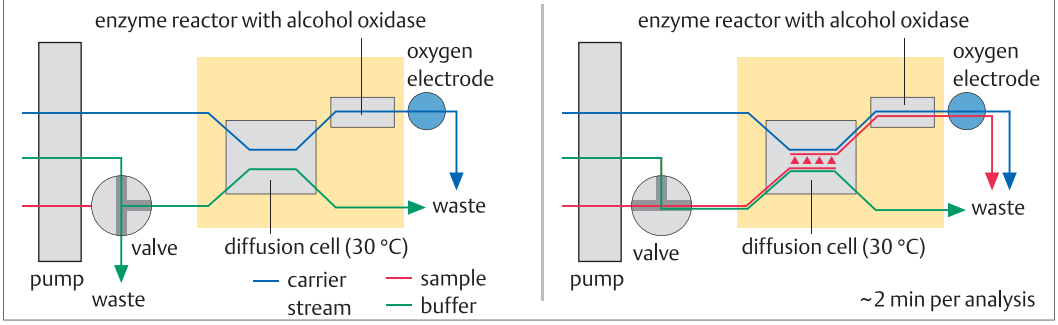
using surface plasmon resonance



once the analyte binds to the antibody, the refractive index in the proximity of the gold layer changes

analyte (antigen)
receptor (anti-analyte antibody)
layer thickness
capillary

Flow injection system for determination of ethanol



Immunoanalysis

General. Although enzymes may allow the rapid and quantitative determination of an analyte in a complex biological matrix such as serum (→256), immunoanalysis has developed into an even better method, since it is more sensitive and much more versatile. The development of radio- and enzyme-linked immunoassays in the early 1970s led to a market which, in 2012, was estimated to exceed 17 billion US$.

Methods. Polyclonal and monoclonal antibodies (→242) bind antigens and haptens with high affinity (k_d usually 10^{-6} to 10^{-8} M). The extent of binding, however, cannot be easily determined. It was therefore a great breakthrough when reporter mechanisms were designed that could detect and quantify the competition between antibodies and antigen binding sites. Usually, homogenous immunoassays, in which no separation step is needed between the binding and the reporter reaction, are distinguished from heterogeneous immunoassays, which include a separation step to remove excess reagent and interfering matrix compounds, but in return are usually more sensitive. Numerous test formats have been designed, allowing the specificity and sensitivity of the test to be adapted to the individual requirements, including the choice of the reporter group (→84). When radioactive isotopes are used (radio immunoassay, RIA), the detection signal for a binding event is formed in a ratio of 1:1. Using enzyme immunoassays (ELISAs), additional amplification of the signal occurs through the enzyme reaction. Well implemented enzyme immunoassays allow for sensitivities down to picomolar or even femtomolar ranges (10^{-12}–10^{-15} M). Horseradish peroxidase or alkaline phosphatase are frequently used as reporter enzymes. If DNA fragments are used as reporter groups, they can be amplified and quantified using real-time PCR (RT-qPCR) (→50). This procedure has permitted extremely sensitive immunoassays, e. g., for prion proteins.

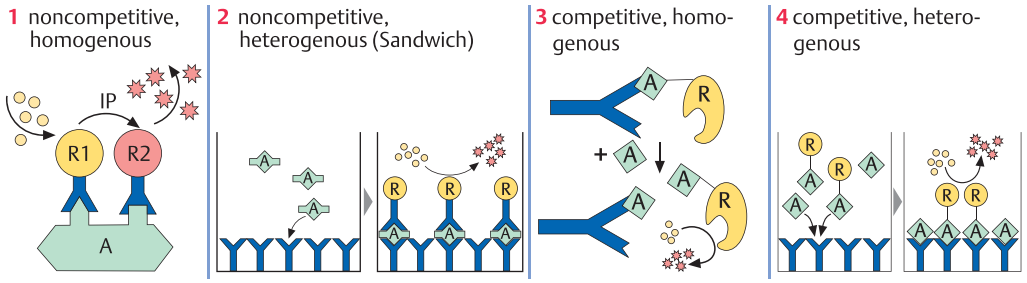
Readout. The results of immunoassays are read through calibration curves. Microtiter plates containing 96 or 384 precoated wells are often used as reaction devices in combination with a microtiterplate reader (a highly parallel photo-, fluoro-, or, preferably, luminometer). In this format, calibration curves can be easily established together with a highly parallel immunoassay.

Analytes. In principle, all substances against which specific antibodies can be raised – either in free form or bound to special hapten carrier molecules – are detectable by immunoassays. Their molecular weight may range from some 10^2 (e. g., drugs, hormones) to about 10^6 Dalton (e. g., multimeric proteins).

Test strips. As with enzyme diagnostics, immunoassays can be performed with test strips, many of which are available in drugstores. For example, pregnancy can be determined by the level of human chorion gonadotropin (HCG), the concentration of which in urine (and in blood) rises strongly after a fertilized egg has implanted in the uterus. Test strips are also used in the hospital for emergency testing ("point-of-care" situations). A typical example is the analysis of troponin T, a protein liberated from damaged heart muscle cells after a heart attack. A drop of blood is placed on the test strip, where it is soaked into the transport zone while erythrocytes are retained. By capillary force, the analyte is transported to the reaction zone, where it elutes specific antibody conjugates, forming a sandwich complex. In a subsequent specific reaction, this sandwich complex binds to a second gold-stained antibody, resulting in a red line. Controls rule out false-positive reactions. The reaction is complete within about 15 min and can be qualitatively monitored by visual inspection or quantified with a CCD camera. A similar immunoassay format is based on the fatty acid binding protein FABP, a 15 kDa protein whose serum concentration rises rapidly after myocardial infarction.

Other examples. Immunoassays have become key diagnostic aids for medical services. With the tremendous growth in the knowledge about the regulation of human metabolism, new diagnostic markers for the early immunodetection of diseases are discovered every year. In food analysis, immunoassays are used to detect the addition of disallowed proteins to food (e. g., of casein to sausages) in a rapid and quantitative manner. The presence of pathogenic microorganisms or toxins in food or water can often be rapidly analyzed using immunoanalysis. Xenobiotics such as herbicides can also be easily detected at high sensitivity (nM) in food or water by these techniques.

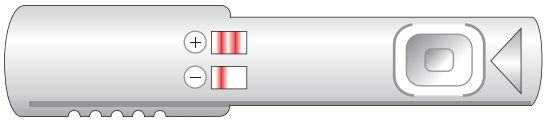
Immunoassay formats



R A reporter-antigen conjugate | A A antigen, hapten | R reporter-antibody conjugate
antibody | enzyme substrate | product | IP = intermediary product

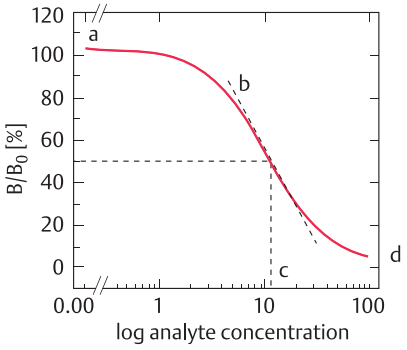
if the reporter is an enzyme, the primary signal is further amplified, enhancing the sensitivity of the assay

Point-of-care kit for heart attack: troponin T



duration of analysis :	15 min
positive:	2 red stripes
negative:	1 red stripe

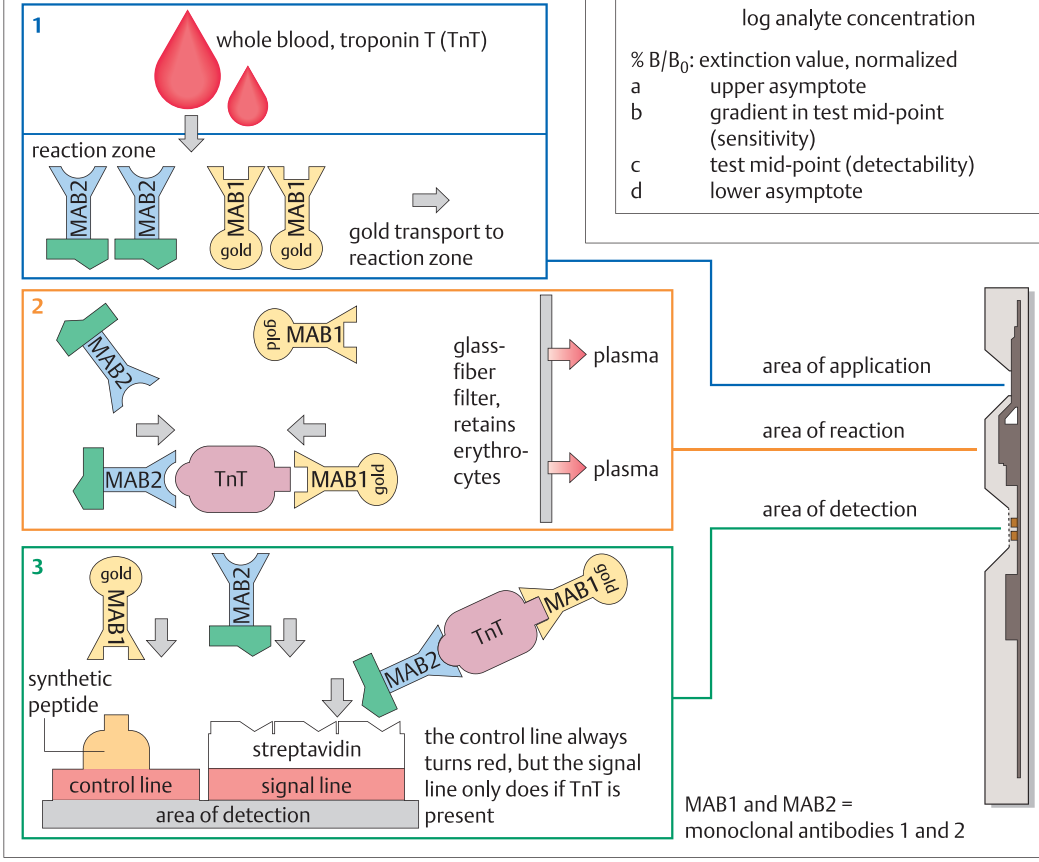
Readout of an ELISA



% B/B_0: extinction value, normalized

a	upper asymptote
b	gradient in test mid-point (sensitivity)
c	test mid-point (detectability)
d	lower asymptote

Mechanism



Glycobiology

General. Sugar chains (glycans) (→32) occur in many biological structures and play an important role in structural biology. The glycans of an organism at large are termed its "glycome." The degree and pattern of glycosylation are important characteristics of many biomolecules, e. g., of antibodies. Glycoproteins (protein part dominates) or proteoglycans (sugar part dominates) are formed by posttranslational modification of proteins in the endoplasmatic reticulum (ER) or in the Golgi. Glycolipids are important structural lipids in plants (monogalactosyl diglyceride) and animals (glycosphingolipids, cerebrosides). Many secondary metabolites of plants (e. g., flavonoids) and microorganisms (e. g., antibiotics) are glycosylated as well. The analysis of glycans may be demanding, as the many chiral hydroxy groups of sugars allow for a large number of regio- and stereoisomers.

Glycosylation patterns. Sugars which often occur in glycans are mannose, galactose, glucose, xylose, N-acetylglucosamine, N-acetylgalactosamine, N-acetylneuraminic acid and N-glycolylneuraminic acid (D-anomers only). They are linked in a regio- and stereoselective manner to an oligosaccharide chain in activated form (e. g., as uridine-diphosphate glucose (UDP-glucose)). In the case of glycoproteins, the peptides are N- or O-glycosylated. N-glycosylation occurs at the free amide nitrogen group of an accessible asparagine within a peptide sequence –Asn-X-Ser/Thr- through transfer of a chain of 14 hexoses bound to the isoprenoid carrier molecule dolichol, which occurs in the ER membrane. This mechanism is conserved in all eukaryotic organisms. O-glycosylation of peptides occurs in the Golgi preferentially on serine or threonine residues. The reaction competes with peptide phosphorylation, leading to more heterogeneous products.

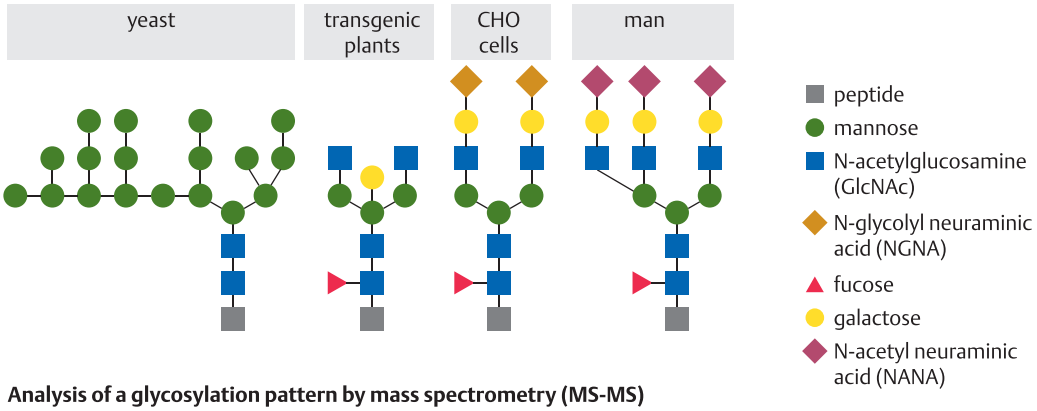
Analytical methods. N-glycans are gently removed from their conjugates by glycosidases. For the removal of O-glycans, ß-elimination can be used. Quantitative analysis of the sugars is done by chromatographic methods after hydrolysis. For glycan sequence analysis, 2D-NMR methods, electron spray (ES) mass spectrometry or more complex mass spectrometric methods such as CAD-MS/MS (CAD = collisionally activated dissociation) are used. Many other methods have been developed. For the analysis of charged sugar chains such as the sialic acids, isoelectric focusing (IEF) or capillary zone electrophoresis (CZE), after treatment with selective glycan-degrading enzymes, are particularly effective. Glycoproteins constitute ~50 % of all proteins; they are analyzed using 2D separation and mass spectrometry (glycoproteomics).

Functions of glycoproteins. Sugar residues on proteins contribute to stability, resistance to proteolysis, viscosity and compartmentation. Glycoproteins are structural components of cell membranes, while proteoglycans may form lubricating molecules (for example mucines). Glycoproteins participate in many cell-cell interactions, e. g., in the immune system (→80), where glycoproteins of the *major histocompatibility complex* (MHC) interact with T-cell-receptors.

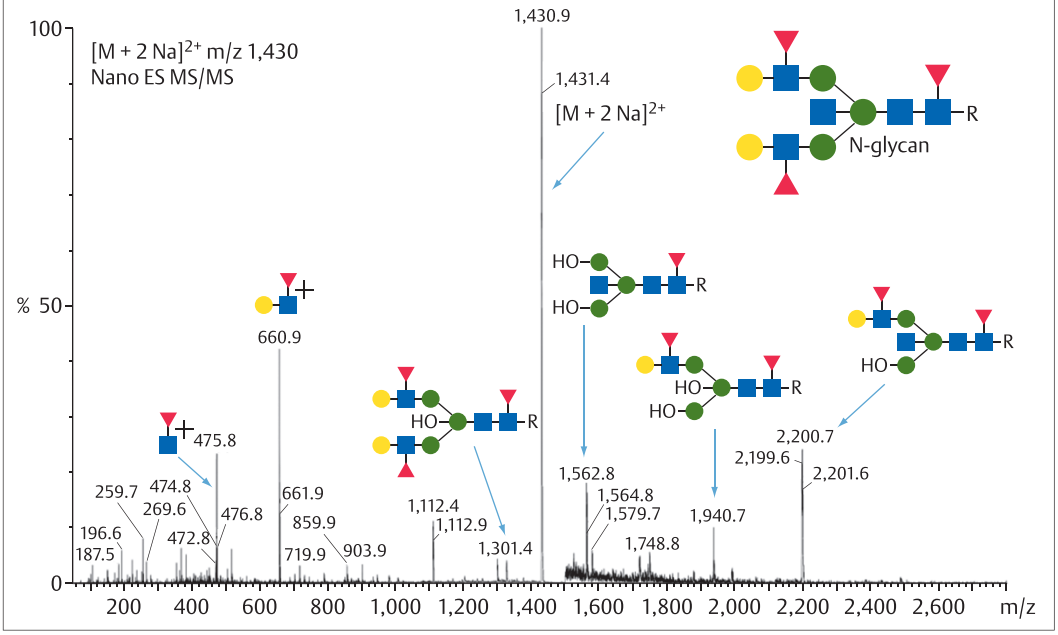
Biotechnological applications. The host systems used for heterologous expression of proteins either do not synthesize glycoproteins at all (*E. coli*) or form glycosylation patterns different from those of humans (baker's yeast, *Schizosaccharomyces*, *Aspergillus*). The precise structure of the glycans may depend on fermentation conditions. "Wrong" glycosylation patterns of recombinant biopharmaceuticals may illicit immune reactions, and it is therefore desirable to adapt their glycosylation pattern to the human original. Thus, services offer expression of recombinant products using either *S. cerevisiae*- or *S. pombe*-strains with controlled posttranslational glycosylation (GlycodExpress® and Aspex®). Using deletions and knockins of suitable glycosyl transferases, it was possible to achieve "humanized" glycosylation patterns. "Glycoengineering" is also applied to many biopharmaceuticals produced in animal cells e. g., CHO-cells (→98). Thus, fucose-free antibodies with largely reduced cytotoxicity for cancer therapy were prepared by deleting the enzyme fucosyl (Fut-8). Human cell cultures are best-suited for the generation of human glycosylation patterns, and several human cell lines have been developed which can be used for the commercial production of antibodies and other biopharmaceuticals with perfect human glycosylation patterns.

Glycosylation and their analysis

Typical glycosylation patterns of organisms



Analysis of a glycosylation pattern by mass spectrometry (MS-MS)



Methods for the manufacture of glycosylated biopharmaceuticals (selection, 2014)

company	process	biological system	applications
Glycode/F	GlycodExpress™	glycoengineered yeast strains	various N-glycosylations
Asahi Glass/J	ASPEX	glycoengineered strains of *Schizosaccharomyces pombe*	
ProBioGen/D	GlymaxX®	removal of fucose in CHO cell processes by a Pseudomonas enzyme	fucose-free antibodies with enhanced ADCC*
Kyowa Hakko Kirin/Lonza, J/CH	Potelligent®	immortalized human cell line	humanized antibodies with enhanced ADCC*
Grucell (Johnson & Johnson)/USA	PER.C6®	human amniocytes	human glycosylation pattern
Gevec/D	CAP®	human amniocytes	human glycosylation patterns
Glycotrope/D	GlycoExpress™	various human cell lines	human glycosylation patterns

*ADCC = Antibody-depended cell mediated cytotoxicity

Animal breeding

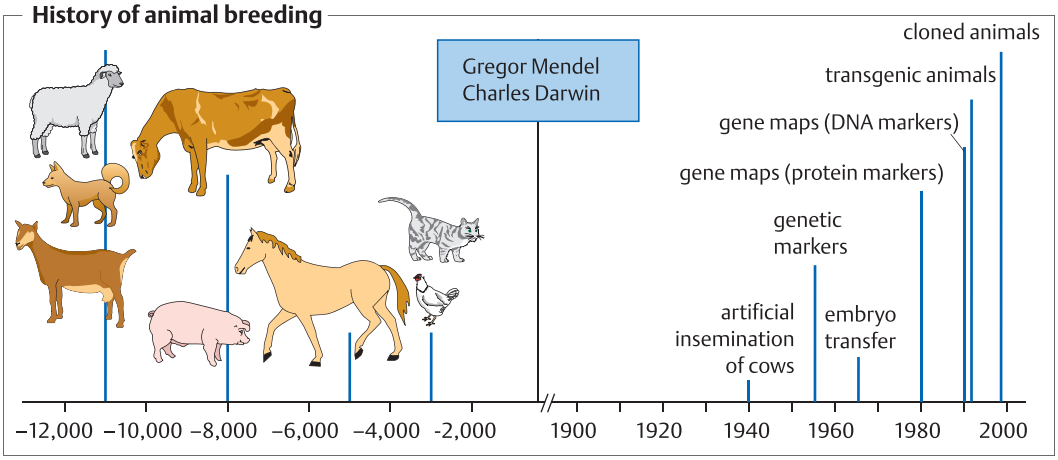
General. Since the "Neolithic revolution" ca. 11,000 years ago, man has domesticated dogs, sheep, and goats, and later (ca. 8000 y ago), also cattle, pigs, and horses. In the beginning, taming and reproduction were the key issues of breeding. More modern targets of animal breeding are the generation of animal products (meat, milk, eggs, wool) with enhanced quantity and quality. For example, modern meat steers may gain > 300 kg y^{-1} of live weight, and dairy cows may yield > 10,000 L y^{-1} of milk – these performance traits values were doubled within just 30 years by breeding programs. Mating of selected animals from prehistoric times has followed phenotypic criteria, whose values depend on genetic and environmental factors. The classical methods of population genetics and biometric analyses are more and more being complemented by modern methods of reproductive biology and gene diagnosis. Examples are 1) artificial insemination; 2) in-vitro fertilization (IVF) and embryo transfer; 3) the preparation of genetic maps (→68) comprising breeding markers; and 4) genotyping markers for performance traits or diseases. Transgenic and cloned animals have, up to now, been mainly used for fundamental and medical research, but there are also examples of, e. g., transgenic silkworms (*Bombyx mori*) used as a "cell factory", and of transgenic sheep, goats and cows producing recombinant proteins in their milk glands (→272).

Artificial insemination (AI). As early as 1729, the Italian doctor Lazzaro Spallanzani described artificial insemination for dog breeding, and even earlier, AI was known for horse breeding in Arabian countries. In 1942, the first insemination station for cattle breeding was established in Germany. AI is not costly and allows the selection of male animals with a high breeding value. From the ejaculate of one bull, 400 portions of sperm containing ca. 20 million sperms each can be obtained, which can be stored at –196 °C. The selection of animals suitable for AI starts with young calves, which have to pass several screening steps based on weight gain, body form, and the milk performance of their mothers. The selected "test bulls" are mated with a number of cows, and the milk and meat performance of their progeny during a test period will decide whether the "bull in waiting" will become a "breeding sire", who, in the end, may replace as many as 1000 conventionally mated bulls for breeding. Cows are inseminated with a portion of thawed sperm by a veterinarian, an insemination technician, or the breeder. In most industrialized countries, > 90 % of cows become pregnant through AI. In pig breeding, ca. 60 % of sows are artificially inseminated.

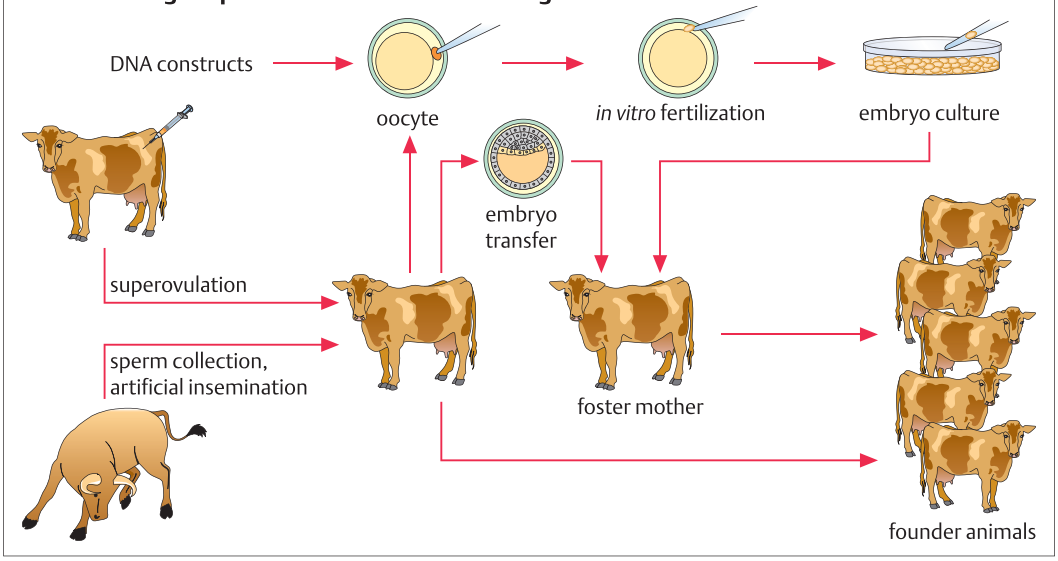
In-vitro fertilization (IVF) and embryo transfer (ET) are mainly applied to increase the number of offspring from high-performing cows. For ET, dams are treated with suitable hormones, resulting in superovulation, which is followed by artificial insemination. This procedure may result in up to 8 embryos suitable for ET, yielding on average 4 live calves after the embryos have been transferred to foster mothers. Although the method is mature for practical application, it is rather complex and expensive and not widely used in agricultural practice. Another well studied method is in-vitro fertilization of egg cells outside the female genital tract, which requires methods for the cultivation and also – for many purposes – the conservation of the resulting embryos. With cows, the eggs are obtained without surgery through ultrasound-based follicle punctation. In other animals, however (sheep, pigs), surgery is required. Sex determination of the embryos can be done by PCR (→50) within 3–6 h (sexing). For interested breeders, sex-sorted embryos are being offered for transfer into foster mothers.

Genetic maps. Genome sequences are now available for many domestic animals such as dog, cat, horse, cattle, chicken, turkey, trout, carp and others. Based on this information and on prior efforts, detailed genetic maps (→68, 268) were established for many domestic animals, especially with concern for those genes that affect performance traits. Gene variants can be analyzed using PCR and RFLP methods. Examples of economic relevance are the ryanodine receptor gene of the pig, in which a mutant largely affects stress tolerance, and variants of the encoding genes that influence milk yield and quality in cows. Most performance traits, e. g., stamina and energy use in racing horses, are influenced by many genes, and their analysis by genetic markers and application for breeding needs further efforts.

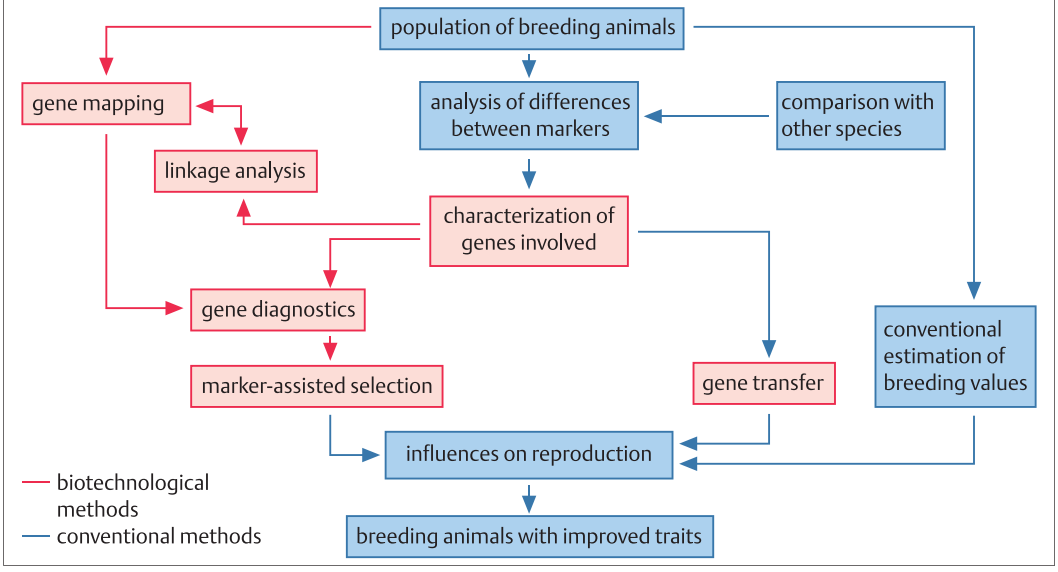
History of animal breeding



Biotechnological procedures in animal breeding



Combination of biotechnological and conventional methods



Embryo transfer, cloned animals

General. This chapter includes methods for superovulation, the cultivation of embryos, and embryo transfer into foster mothers. Transgenic embryos and the production of cloned animals is also addressed.

Embryonic development in mammals (→78)**.** The egg of most mammals is released from the ovary during metaphase of the second meiosis (oocyte II). If fertilization by a sperm ensues, meiosis continues and the second polar body is released. The two haploid prenuclei, originating from the oocyte and the sperm, fuse to form the diploid nucleus of the zygote, and the first cell division ensues. During further divisions, up to the morula stage, the size of the nucleus continuously decreases. The morula stage sees the first cell differentiation, leading to the blastocyst, which implants, after the zona pellucida has been dissolved, into the mucousa of the uterus, resulting in an embryo.

Superovulation and embryo cultivation. With most mammalian species, the probability that excess oocytes are ovulated (superovulation) can be increased by hormone treatment of the mother. Often it is also possible to obtain egg cells, fertilize them in vitro (IVF), and develop them ex vivo up to the embryo stage, using suitable nutrient media. For economic reasons, most experiments have been done with cattle and sheep embryos. Both can be conserved indefinitely at –196 °C (cryopreservation).

Embryo transfer (ET) and embryo splitting. ET is the transfer of foreign embryos into a foster mother of the same species. The embryos may originate from donor animals that were superovulated and artificially inseminated. This is different from embryo splitting, in which blastomeres are isolated by microsurgery from one morula; each of a group of blastomeres is cultivated in vitro to the stage of a blastocyst. After transfer into a foster mother, these blastomeres may develop into genetically identical animals. Using this procedure, usually 2, and sometimes 6–20 transferable embryos can be obtained from a single cow, of which ca. 50 % will develop into healthy calves. This method is practiced in agricultural breeding.

Transgenic embryos. During the oocyte or blastocyst stage of embryonic development, gene constructs can be introduced by microinjection into the prenuclei or into embryonic stem cells. By this method, new genetic material can be introduced into the recipient embryo. The transgenic embryos originating from this manipulation can be transferred into foster mothers, who eventually give birth to transgenic animals. Pioneering studies in this area were done with mice, since they are important laboratory animals for both fundamental and applied research. The procedure is also applied to farm animals and was a key issue for "gene farming" (→272).

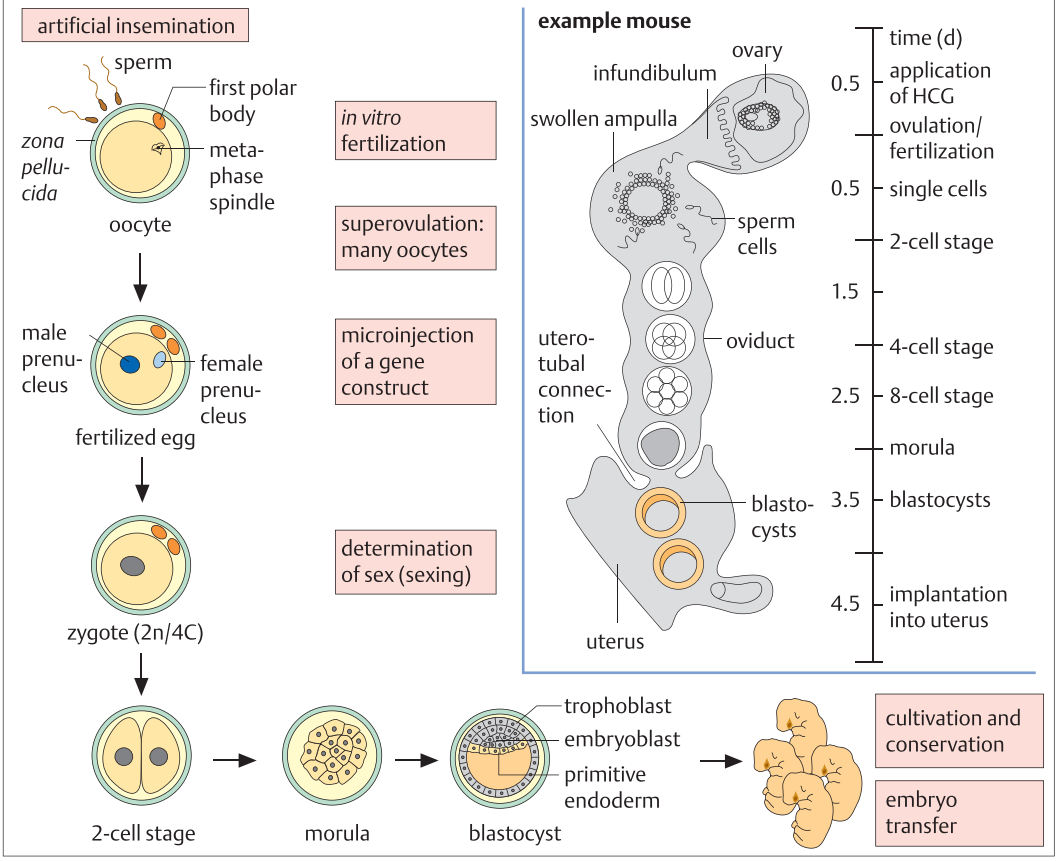
Cloned animals. Asexual reproduction leading to identical clones is widely distributed in unicellular organisms, plants, and lower animals. In higher animals, genetically identical clones are quite rare, e. g., the frequency of homozygous twins in humans is only 0.3 %. For experimentally generating clones of higher animals, an egg (haploid chromosome set) is taken from a female donor, and its nucleus is removed with a micropipette. In somatic cells of the same animal species (e. g., obtained from a cell culture of udder epithelium), the G_o phase of the cell cycle (when no cell division takes place) is induced, and the cell, containing a nucleus with a diploid chromosome set, is fused to the enucleated egg cell. The resulting diploid cells are developed to the embryonic stage either in cell culture or in the oviduct of a sterile female, to be transferred into a foster mother. In 1997, for the first time an animal genetically identical to its mother was created based on differentiated somatic cells (Dolly the sheep). It was the only successfully developed lamb in an experiment using 277 enucleated eggs and a total of 27 embryos that were derived from them. In spite of these experimental drawbacks, the method was further improved and successfully transferred to dozens of animal species such as mice, goats, pigs, and cattle. Some experimental progress has been made applying this technology for the generation of transgenic, monoclonal herds, e. g., to produce therapeutic proteins in their milk (gene farming), for the propagation of high breeds such as racing horses, and also for the preservation of endangered species.

Hormones controlling the female cycle of mammals

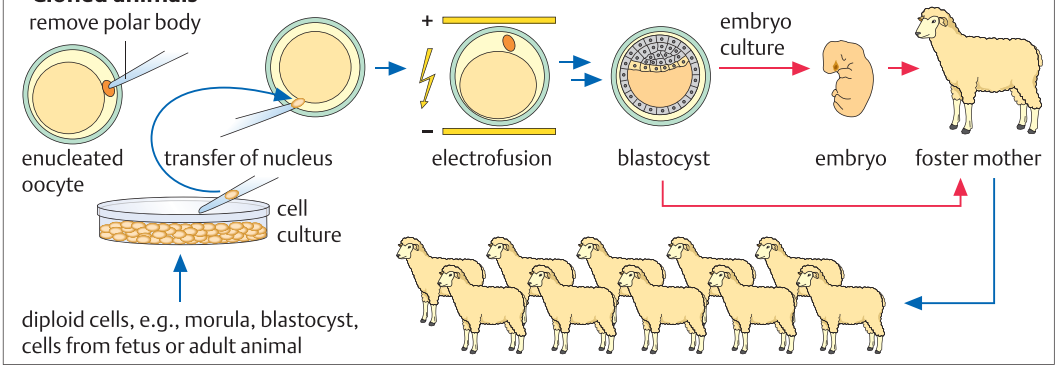
bioactive agent	isolation	effect
prostaglandin PGF24	chemical synthesis	degradation of *corpus luteum*; release of sexual heat
follicle-stimulating hormone (FSH)	pituitary of pigs or recombinant protein	follicle stimulation leads to superovulation
pregnant mare serum gonadotropin (PMSG)*	from serum of pregnant mares	follicle stimulation leads to brunt and superovulation
gonadotropin-releasing hormone (GnRH)	chemical synthesis	induces ovulation

*synonym: eCG = equine choryogonadotropin

Embryonic development of a mammal and biotechnology



Cloned animals



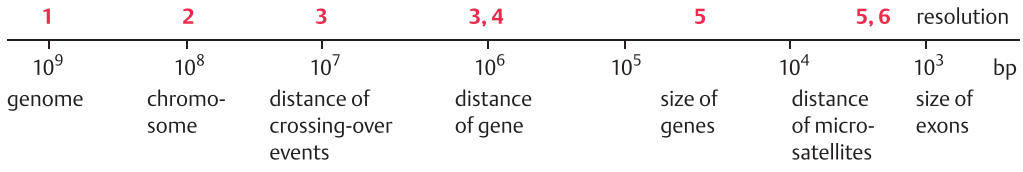
Gene maps

General. During the past 100 years, animal breeding has been carried out by selection and mating. Statistical methods were developed to analyze genetic and environmental influences. For the important domestic animals (horse, cattle, pig, sheep, goat, chicken, dog, cat), many of whose genomes have now been sequenced, increasingly precise genetic maps were developed, which are based on the linkage of inherited traits or markers. Physical maps, in contrast, indicate the position of genes on the DNA of individual chromosomes (→72). The genomes of domestic animals are usually on the order of 3 Gbp and of similar complexity as the human genome. So far, genome sequences have been completed for the dog and the chicken (2005), the horse (2007), followed by the bovine genome (2009) and many others. Moreover, for many economically important domestic animals, genome-wide coupling maps do exist, which comprise for each species several thousands of microsatellite loci as markers and several hundred genes whose functions have been assigned. Using these data, the allelic variants of these genes can be typed, using PCR methods. Finally, the genetic variants can be associated with performance traits and used for breeding.

Genetic improvements. The assignment of desired traits (e. g., high milk production) to the genes resembles the task of assembling a puzzle made up of tens of thousands of pieces of high similarity but diverse origin (parents and progeny). To simplify genetic analysis, environmental factors involved in the trait development first had to be standardized. In breeding practice, the influence of genetic vs. environmental factors on the variation of trait values can be estimated by complex statistical methods (e. g., BLUP, best linear unbiased prediction). With methods such as artificial insemination and embryo transfer (→264, 266), groups of domestic animals can be obtained in which half the gene pool originates from one parent, allowing a more precise analysis of the genetic origin of traits in the progeny. However, none of these methods allows the effect of individual gene polymorphisms on complex traits to be analyzed in a straightforward manner.

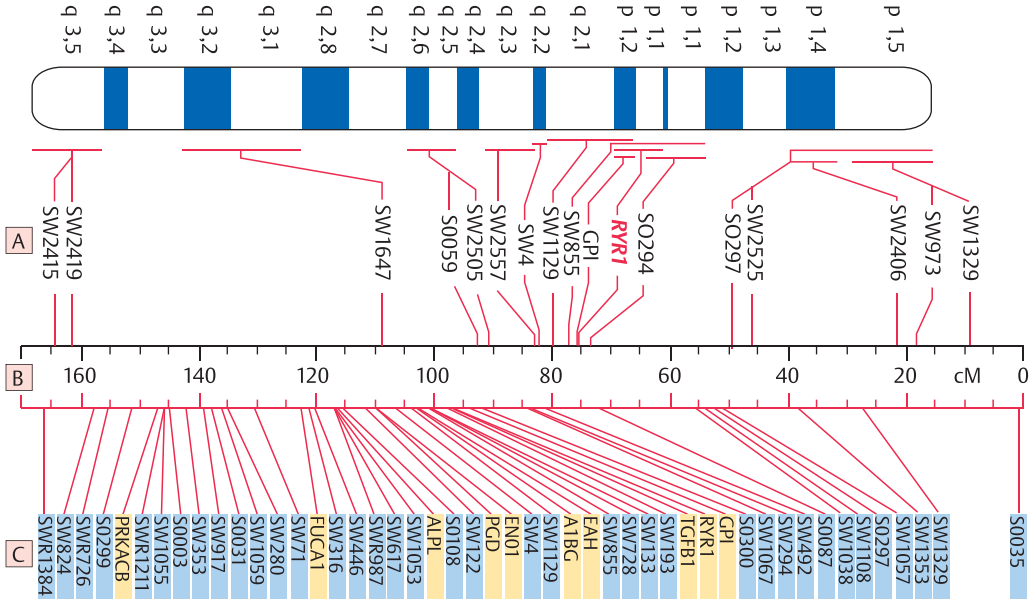
Gene maps and genome sequencing. Since the genomes of most domestic animals comprise ca. 3 Gbp, and direct sequencing of DNA is limited to a length of ca. 600 bp per assay, complex procedures are necessary to localize individual genes. Of key importance is the identification of DNA marker sequences that are polymorphic in the parental DNA (i. e., they are composed of different alleles) and trace the inheritance of allelic markers in offspring generations (→72). Microsatellites are the most important markers used for this purpose. They have variable numbers of tandem repeats (VNTR) and are frequently found in mammalian genomes (about 50 000–100 000 microsatellite loci per genome) (→72, 298). A second method for analyzing the positions of DNA markers is RFLP analysis (restriction fragment length polymorphism). To this end, parental and progeny DNA are digested with endonuclease, and the resulting patterns of fragments are compared by gel electrophoresis. This method leads to maps for polymorphisms that correlate, in favorable cases, with values of performance traits, rendering them useful for breeding experiments. Once the location on the genome of a polymorphism important for breeding has been identified, the locus can be amplified by PCR, and sequence analysis can be carried out using the PCR product. However, care must be taken that enough information is available about the intron–exon structure of the pertinent gene, preferentially through analysis of the transcribed mRNA or the cDNA derived from it. For economically important domestic animals such as chicken, cattle, and pig, gene maps are already available which have been obtained by observing the inheritance of traits, genes, and DNA markers. Linkage analysis provides further information about the relative positions of genes and markers on chromosomal DNA, since recombination events during meiotic crossing over are recorded by this method. Within the international "1000 bull genomes project", > 200 animals of different races have been sequenced in an attempt to identify, by genome-wide association studies, those alleles which are relevant for breeding. A new dimension for breeding genetics is the finding that epigenetics play an important role, e. g., in breeding for stress or disease resistance. Epigenome-wide association analysis will be needed to detect which epigenetic traits are associated with undesired traits in breeding.

Methods for gene mapping



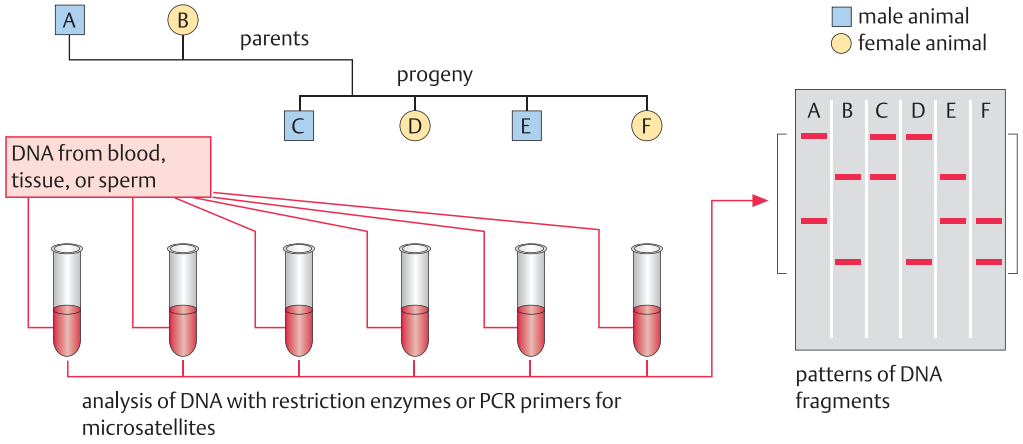
1 crossing, linkage analysis
2 cell fusion (somatic cell hybrids), *in situ* hybridization
3 linkage analysis
4 limit to clonable gene fragments, restriction fragment length polymorphism (RFLP)
5 PCR analysis
6 gene sequencing

Gene map of chromosome 6 of the pig



A physical map and probes used for *in-situ* hybridization
B genetic map (in cM)
C blue: microsatellites (marker loci), yellow: gene segments, e.g., *RYR1*, a receptor protein assiciated to the phenotype "stress susceptibility"

RFLP and microsatellite marker



Transgenic animals

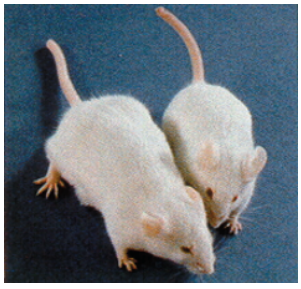
General. Transgenic animals, which express foreign genes (knock-in) or lack expression of indigenous genes (knockout) (→64), play an important role in fundamental research, as animal models for human diseases, and for animal breeding. Due to its physiological closeness to man and its simple breeding, the mouse (*Mus musculus*, adjective: murine) is the most important animal model.

Transgenic animals. To minimize side effects due to heterozygous genotypes, transgenic animals are preferentially derived from parental inbred lines. In mice, inbreeding between brothers and sisters for 7–10 generations results in largely homozygous populations. For the transfer of foreign genetic material, microinjection of gene constructs into the prenuclei of embryonic stem cells (ES) is mostly used (→78). Although this method is fast, only a few embryos survive this treatment. Suitable vectors usually contain the gene construct in an intron–exon structure, a promoter, and a polyA signal sequence. Recombination of this sequence into the recipient animal genome often occurs as several copies at several positions simultaneously. Since this is undesired for knockout experiments, transfection of embryonic stem cells in vitro is preferred. Insertion vectors that contain DNA stretches homologous to the gene to be replaced, but lack sequences that would be important in coding for a functional gene, are used. Recombination is initiated by inducing double strand breakage or crossing over and results in inactivation of the functional gene. Selection markers or gene traps, where a reporter gene is cloned into an exon or regulatory sequences of the target gene, make it easier to control the experiment. Using such gene-targeting protocols, recipient cells can be selected which have only one gene copy integrated into a single well defined position in the DNA sequence. To silence genes, various kinds of antisense or interfering RNA constructs are also being intensely applied ("genome editing," →64). Progeny containing the newly recombined or silenced genes in some of their sexual cells (germline chimeric animals) are used for further breeding of homozygous transgenic animals (founder animals).

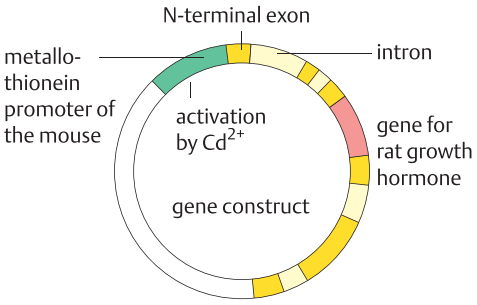
Transgenic mice. Superovulation is induced in inbred females by injecting gonadotropin (HCG) intraperitoneally, then they are made pregnant by mating. From their oviduct, multicell stages, morula, or blastocysts can be obtained, depending on the time of preparation. These cells can be transformed by transfection or microinjection. In the transfection protocol, pluripotent embryonic stem cells (ES cells) (→78) are prepared from blastocysts and propagated in tissue culture. They can be transformed using suitable DNA vectors and reinjected into blastocysts, resulting in transgenic embryos. In the microinjection protocol, a suitable vector with foreign DNA is injected into the larger male prenucleus (which becomes visible in egg cells after fertilization). In both methods, the transformed blastocyst or egg cell is implanted into the uterus of a foster mother which was made pseudopregnant by mating with a vasectomized male mouse.

Applications. In 1982 for the first time, a gene construct for rat growth hormone (→224) was microinjected into the prenucleus of a fertilized mouse egg, which developed in a foster mother into a transgenic "super-mouse". Using this type of technique, particular genes can be analyzed for their involvement in genetic diseases. Thus, e. g., by knockout experiments, experimental evidence may accrue as to whether this gene is involved in the pathogenic state. In the "oncomouse", the activated v-Ha-ras oncogen was coupled to an embryonic promoter which, upon injury of the epidermis, led to the formation of skin tumors; oncomice can thus be used for dermal testing of mutagens. Transgenic mice with mutations in the β-amyloid precursor protein (APP) serve as an animal model for Alzheimer's disease, the SCID (severe combined immunodeficiency) mouse with a genetic immunodeficiency as a model for immune diseases. By the end of 2013, about 50,000 mouse genotypes with known phenotypic annotations were entered into the *Mouse Genome Informatics* (MGI) database, and mouse models were available for a total of 1,274 human diseases. Since genome sequences for many mouse strains are available, analysis of transgenic mice is also a valuable tool for functional analysis of the human genome. Apart from mice, rats are important experimental animals, and transgenic rats are also commercially available. There are also studies on, e. g, SCID swine, as the pig is an excellent biomedical model for humans.

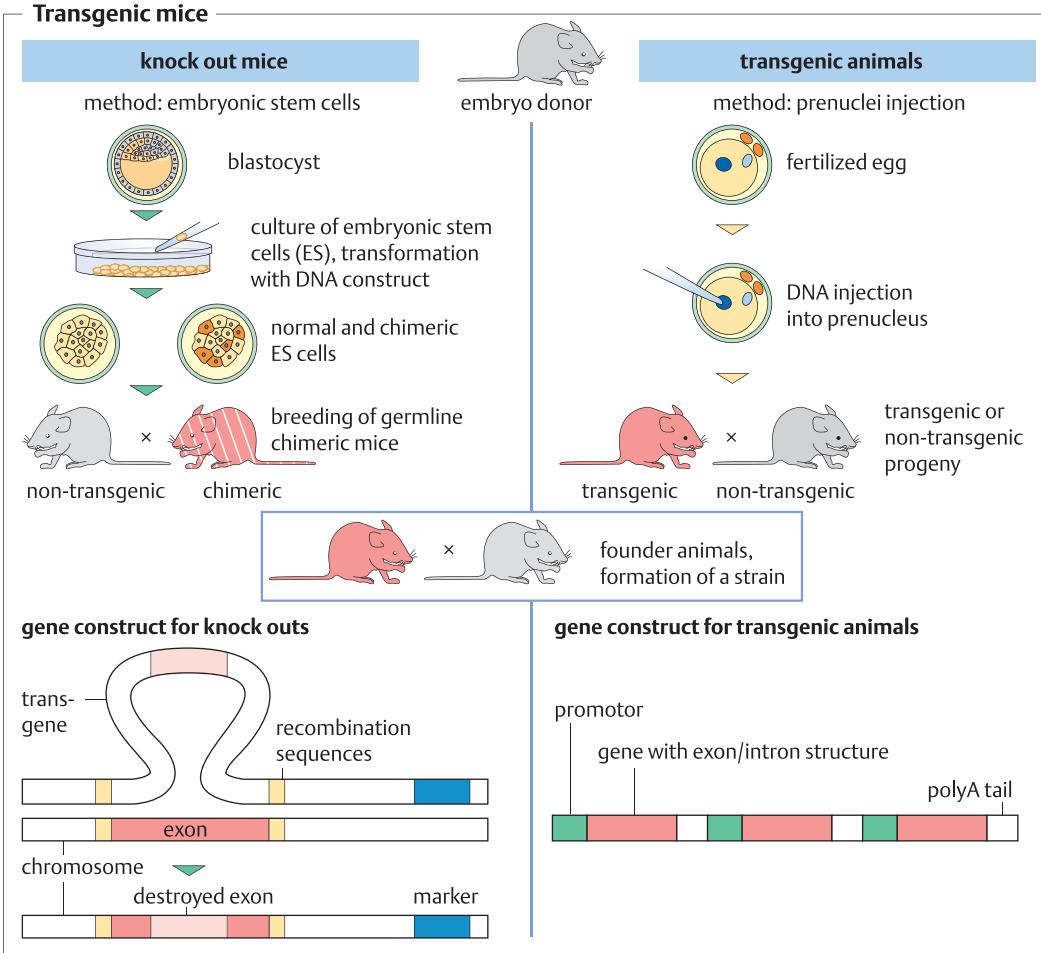
"Supermouse"



after microinjection of a gene construct containing the rat somatotropin gene into the prenucleus of a fertilized mouse egg (right side: control)



Transgenic mice



Mouse strains* for clinical research

changes/defects	applications
conventional breeding	nude or hairless mice for skin compatibility tests
expression of rat somatotropin	"supermouse"
deficient immune system	SCID mouse: immunodeficiency
defective p53 tumor-suppressor protein	oncomouse: carcinogenesis
angiotensinogen defective	high blood pressure mouse
CTFR defect	cystic fibrosis mouse (for gene therapy experiments)

*as of 2013, >1200 types of transgenic mice had been reported as models for human diseases

Breeding, gene pharming and xenotransplantation

General. Recombinant DNA technologies are widely used for research on improving animal production. Examples include the breeding of transgenic cattle which produce lean meat and lactose- or allergen-free milk, transgenic sheep with enhanced biosynthesis of cysteine, leading to a higher wool production, transgenic silkworms which produce specialty silks (→156), and transgenic pigs which express recombinant phytases (→192), enabling the hydrolysis of phytic acid-bound phosphorous. Transgenic dogs, chicken, salmons and racing horses are also being studied. Due to the demanding technology and public controversy (→336), transgenic animals have so far been used in some areas only. For example, goats, sheep, and cattle have been engineered to produce pharmaceutically valuable proteins in their milk (gene farming). Similar to related techniques in transgenic plants, yields are often surprisingly high and competitive with technical procedures using recombinant animals or animal or microbial cells in a bioreactor. Especially for heart transplantation, transgenic pigs are being considered for the provision of replacement organs for humans (xenotransplantation).

Breeding of transgenic animals was originally focused on increasing growth by gene transfers resulting in, e. g., increased endogenous production of growth hormone (somatotropin). Recently, other goals, such as enhanced resistance to disease and stress, as well as improvement of meat or egg quality, are being addressed. In mammals, gene transfer was done by microinjection into prenuclei of fertilized egg cells and embryo transfer (→264, 266). Chickens are transformed using recombinant retroviruses or by fertilization with recombinant sperm. Transgenic fish can be obtained by electroporation of egg cells with DNA. An examples for the application of these techniques is improvement of cold resistance of fish by cloning in an antifreeze protein.

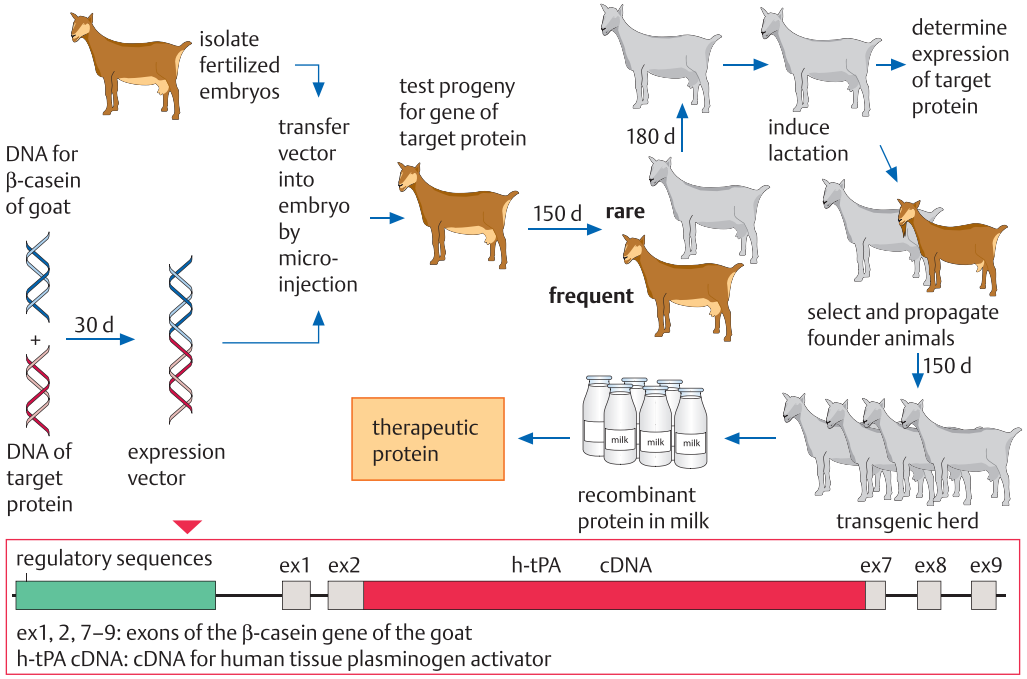
Gene pharming. Biomedically important proteins can be secreted into the milk of transgenic animals. The desired gene is often cloned behind the β- or αS1-casein promoter, and a suitable gene construct is injected into the prenucleus of a fertilized egg cell, resulting in a transgenic embryo. Although the success rate of this technology is still modest (often < 0.1 % of the treated egg cells develop into embryos), transgenic animals have already been developed that produce recombinant proteins such as α_1-antitrypsin (→232), tPA (→230), urokinase, IGF-1, IL-2 (→236), lactoferrin, or human serum albumen (→226) in yields up to 35 g L^{-1} milk. The products can be isolated from milk or used directly in milk. Since a high-performance dairy cow produces up to 10,000 L milk per year, a single transgenic dairy cow would produce the quantity of factor VIII (→228) required for the entire USA (120 g). Human antithrombin (ATryn) (→230), a plasma protein with anticoagulant properties, is produced in goats and registered for human use both by the FDA and EMEA.

Xenotransplantation. As of 2014, > 120,000 US citizens wait for organ transplants; among them, ~100.000 wait for a new kidney, and ~ 3.000 wait for a new heart to be transplanted, although only ca. 2,000 hearts per year are available. Against this background, transgenic animals are being studied as organ donors, the most suitable animal being the pig: its organs are similar to human organs in size, anatomy, and physiology. In xenotransplantation, the key issue is organ rejection by immunoreaction (→80). In this context, we can distinguish 1) hyperacute immunorejection (seconds to minutes), based on rapid activation of the complement system in the recipient; 2) acute immunorejection (days), based on the reaction of T cells; and 3) chronic immunorejection (up to several years), whose precise mechanism is unknown. The prime goal when transplanting organs from other species is to prevent hyperacute immunorejection. To this end, transgenic pigs were obtained whose complement cascade is replaced by human factors. A key improvement was the cloning of hCD55 (decay accelerating factor, DAF). When hearts of transgenic pigs containing this factor in their complement system were transplanted into primates, the recipients survived for ca. 40 d, whereas controls died within a few minutes. The knock-out of α1,3-galactosyltransferase in pigs is another target, since α-linked terminal galactose residues occur only on porcine organs and tissues and, in humans, give rise to the formation of anti-α1,3-Gal antibodies.

Improved performance in transgenic animals

	effect	target protein or coding gene
pig	malign hyperthermic syndrome Ca metabolism in muscle	ryanodin receptor gene (*RYR1*) on chromosome 6: cys ➔ arg
cattle	κ-casein: milk quality	κ-casein gene
salmon	*antifreeze* protein (AFP) for enhanced temperature tolerance	expression of an AFP-encoding gene from the winter flounder

Manufacture of a pharmaceutical product with transgenic goats



Espression of therapeutic proteins in the milk of transgenic animals

gene construct	recombinant protein	quantity expressed [mg/L]	transgenic animal
WAP	tPA (cDNA)	0 – 50	goat
WAP	protein C (cDNA)	1,000	pig
BLG	α_1-antitrypsin (genomic DNA)	35,000	sheep
BLG	human serum albumen	10,000	mouse
β-casein of goat	tPA (cDNA)	3,000	goat
αS1 casein of cow	urokinase (genomic DNA)	2,000	mouse
αS1 casein of cow	insulin growth factor (IGF-1) (cDNA)	10,000	rabbit

WAP = whey acidic protein BLG = bovine β-lactoglobulin

Comparison of productivity

protein	market in USA (kg/y)	from human plasma (L)	number of dairy cows (milk production = 10,000 L/y)
factor VIII	0.120	1,200,000	1.2
protein C	100	20,000,000	100
fibrinogen	200	500,000	20
α_1-antitrypsin	800	4,000,000	80
serum albumen	100,000	2,000, 000	5 000

Plant breeding

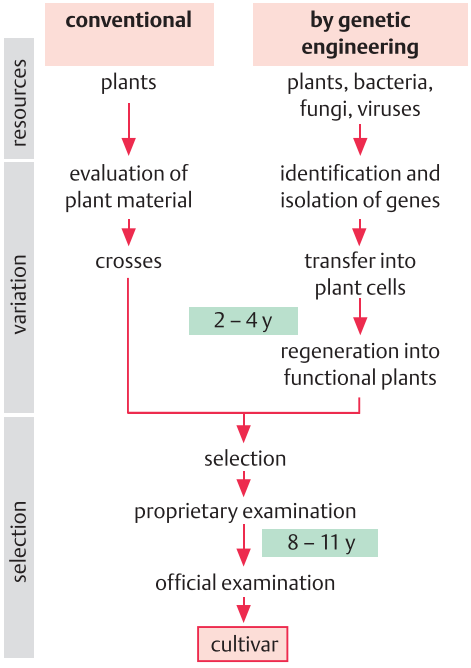
General. About 11,000 years ago, man began to cultivate plants. As the result of this long breeding process, today's cultivated plants yield much more biomass, fruits, and seeds than their wild ancestors. The nutrition of man and his domestic animals relies to a large extent on these cultivated plants. Even today, however, about 25 % of the 7.2 billion humans on earth (2014) are inadequately nourished. Since the human population is expected to grow to about 12 billion by the year 2100, an increase in plant productivity will be a key requirement for the future.

Plant breeding. The result of plant breeding is called a cultivar: a plant line with specifications that are typical for this variety and that are inherited upon propagation. A variety arises from crossing and selection. Depending on the type of propagation, pollen source, and the plant's genetic structure and composition, lines, populations, synthetics, clones, and hybrids can be distinguished. Genetically homogenous varieties are obtained with self-fertilizing plants such as wheat, rice, barley, and sugar cane. Other flowering plants such as maize, potato, soybean, and sugar beets, however, are outbreeding and thus highly heterozygous. If they can be propagated vegetatively, as is true for e. g., potatoes and sugar beets, synthetic or clonal varieties of a narrowed genotype can be obtained. Outbreeding plants can be bred into highly heterozygous but completely homogenous hybrid varieties through enforced self fertilization (inbreeding). In the breeding of some plants such as maize, the male inflorescences are removed for this purpose. If male and female organs are combined in one flower, however, this type of manipulation is difficult. A solution to this problem is the use of male-sterile parental lines, which are obtained by either of two methods: from varieties with cytoplasmic male sterility (CMS, coded by the mitochondrial genome) or by using self-incompatible (SI) lines, a widely distributed mechanism for the prevention of self-pollination. Heterozygous individuals, however, are often stronger than homozygous ones, presumably because their heteroallelic gene products are inactivated with less probability or else show a wider functional range. This property (heterosis) is often used for backcrossings, similar to the methods used for antibiotic-producing microorganisms. Similar protocols are used in the horticulture industry, which has developed more than 11,000 varieties of crops, with a business value of 12 billion US-$ in the US alone (2009).

Forestry. In the early history of man, the uncontrolled cutting of forests led to erosion and denudation of large areas of land. This process continues to occurs today in tropical rainforests. Only since ca. 1800 has a sustainable forest economy been established in Central and Northern Europe (e. g., pines must reach an age of ca. 100 y before being cut, oaks ca. 300 y). Wood (annual production ca. 7×10^{10} t) is a valuable renewable resource and will probably be used to a much greater extent in the future for the production of chemical base materials, including raw materials for fermentation. Today, it is predominantly used in manufacturing lumber and chipboard, and for the paper and cellulose industries (→184).

Biotechnological procedures. Generating and screening novel genotypes can be accomplished by several methods. Chemical mutagens, radiation and transposons have been used to generate mutants. Induced chromosome multiplication (polyploidy) or chromosome elimination has also been applied. Male sterility, which is important for the breeding of homozygous hybrids, can be achieved by genetic engineering techniques, e. g., through the expression of a highly active RNase from *Bacillus amyloliquefaciens* under a pollen-specific promoter, leading to pollen inactivation. This process can be modulated through expression of a restorer gene, which inhibits RNase activity. The preparation of callus, meristem, protoplasts, or haploid cultures (→276), which can often be regenerated into intact dicot or monocot plants has greatly increased the speed of plant breeding efforts. During regeneration, in particular from callus cultures, somaclonal variations may occur which can be selected or removed. Transgenic plants can express resistance factors against viruses, fungi, bacteria, herbicides, or insecticides (→282), but can also be engineered to express high-value products (→284). The progress in the total sequencing of plant genomes has led to a much higher precision in the location of markers within plant gene maps and has contributed to establishing new target-oriented strategies in plant breeding.

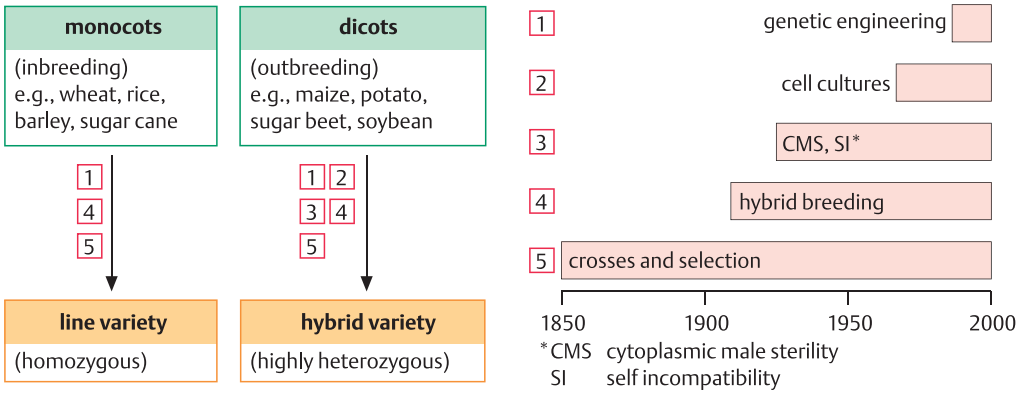
Methods of plant breeding



Production of agricultural plants (2000)

product/plant		world production [10^6 tons]
sucrose	*Saccharum officinarum*	1 260
maize	*Zea mays*	594
rice	*Oryza sativa*	594
wheat	*Triticum aestivum*	584
flax	*Linum usitatissimum*	503
potato	*Solanum tuberosum*	321
beet sugar	*Beta vulgaris*	246
cassava	*Manihot esculenta*	175
soybean	*Glycine max*	161
sweet potato	*Ipomoea batatas*	142
barley	*Hordeum vulgare*	132
tomato	*Lycopersicon esculentum*	98
wine	*Vitis vinifera*	64
sorghum	Andropogonoideae	58

Technological advances



Size of plant genomes

	species	number of chromosomes	ploidy	genome 1 000 Mbp
monocots				
barley	*Hordeum vulgare*	7	2	4.8
rice	*Oryza sativa*	12	2	0.42
wheat	*Triticum aestivum*	7	6	16
dicots				
maize	*Zea mays*	10	2	2.5
Arabidopsis	*Arabidopsis thaliana*	5	2	0.1
rapeseed	*Brassica napus*	19	2	1.23
tomato	*Lycopersicon exculentum*	12	2	1
tobacco	*Nicotiana tabacum*	12	4	4.4
potato	*Solanum tuberosum*	12	4	1.8

Plant tissue surface culture

General. In the past 40 years it has become possible to propagate tissues and cells of plant organs (roots, leaves, etc.) as organ cultures. Through treatment with plant hormones, such cultures can often be regenerated into intact, fertile plants. This method is widely applied in fundamental research, but is also used to 1) produce large numbers of plants from a stock plant by micropropagation; 2) breed plants with improved properties (in-vitro selection); 3) propagate ornamental and garden plants in a virus-free manner (→6); 4) produce transgenic plants; and 5) conserve plant species menaced by extinction as regenerable cell cultures (germ plasms). Plant tissue culture can also be used for preparing secondary plant products at a technical scale.

Methods. In a procedure termed somatic embryogenesis, tissue from the desired organ of a plant that has been grown from a sterile seed (an explantate) is transferred under aseptic conditions to a growth medium. Since most cells and tissues are heterotrophic (→12), growth requires a carbon source such as glucose or saccharose and nitrate as a nitrogen source. The medium also contains vitamins, trace elements, plant cytokinins such as kinetin or zeatin, and growth factors such as 3-indolyl acetic acid, 2,4-dichlorophenoxyacetic acid, or abscisic acid. For hard-to-cultivate cells, more complex media have been used. The cultures are usually grown in chambers under sterile conditions where light, humidity, and temperature are controlled (climate chambers). Depending on the cultivation conditions and starting material, callus or suspension cultures, or meristem or haploid cultures are obtained.

Callus cultures. A wounded tissue that grows in an uncontrolled manner is termed a callus. In plants, it originates from the planes of section of the explantates. Calluses can be cultured on the surface of agar plates and have been used as the starting material for the production of undifferentiated, omnipotent plant cells. Completely differentiated plants can be regenerated by treating callus cultures with plant hormones, provided that they have not become too old.

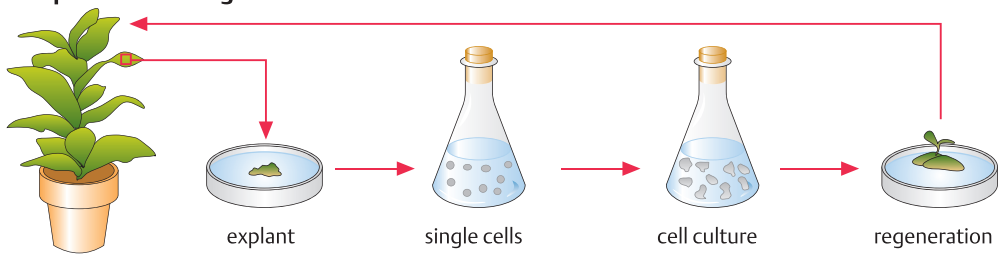
Suspension cultures. Similar to microorganisms or animal cells, plant cells which have been obtained, e. g., by maceration of a plant organ, can be propagated in sterile liquid nutrient media (→278). Compared to callus cultures, suspension cultures contain a higher variation in cell types and are more difficult to maintain in a standardized manner.

Meristem cultures. Meristems are the embryonic cells of plants, which can divide indefinitely. They can be isolated under sterile conditions as leafless vegetation cone from shoots, roots, or axillary shoots and will grow as callus or suspension cultures. A meristem culture is uniquely suited for obtaining and mass propagating specific pathogen-free plants. To this end, meristemic cells are subjected to a short heat treatment at 40 °C; high yields of pathogen- and virus-free cultures are obtained, probably through the formation of heat-shock proteins. From such cultures, pathogen- and virus-free plantlings can be regenerated. These are, however, not resistant and are subject to new infections. Meristem cultures of grapevines, strawberries, banana, sugarcane, or potatoes and of ornamental plants such as carnations, lilies, chrysanthemums, and orchids have revolutionized gardening practices. The world market for virus-free seed or plantlings is estimated as several billion US$ per year.

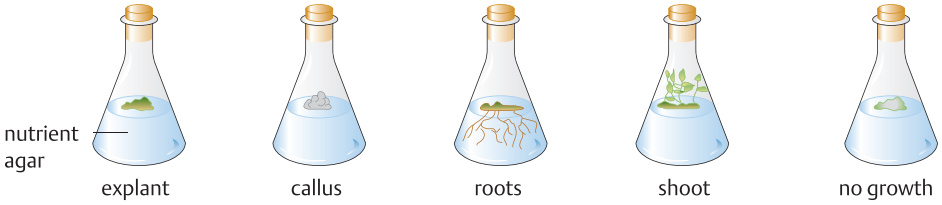
Haploid cultures are cell cultures of plant sexual organs, in particular, the microspores. After propagation in surface cultures, they can be regenerated either into sterile haploid plants with only one set of chromosomes or, in the presence of the mitotic poison colchicine or after protoplast fusion (→278), into homozygous diploid plants (doubled haploids, DH). Such plants are of great help to the breeder since they pass on a set of constant traits to their progeny. Most economically relevant traits are caused by the interplay of several genes, and the effect of variations in a single gene are small. Doubled haploids breed true and can be propagated in large numbers. As a consequence, they are very useful to map quantitative trait loci (QTL). Haploid cultures are being used in breeding potatoes, barley, rapeseed, tobacco, and some medicinal plants.

Somaclonal variation. Although clonal fidelity is high in conventional micropropagation, chromosomal instability of cultured plant cells may lead to genotypic or phenotypic changes upon regeneration of whole plantlets and are used to recover novel genotypes useful for crop improvement from cell cultures.

Preparation and regeneration of a meristem culture



Influence of phytohormones



auxins	–	3.00 mg/L	3.00 mg/L	0.03 mg/L	–
cytokinins	–	0.2 mg/L	0.02 mg/L	1.00 mg/L	0.2 mg/L

	structure	function*
auxins	3-indolyl acetic acid, (2,4-dichlorophenoxy) acetic acid (a synthetic auxin) and others	induce longitudinal growth, higher concentrations inhibit formation of roots, cell division
cytokinins	kinetin, 6-benzylaminopurine, gibberellins, and others	stimulate callus formation, cell growth, and longitudinal growth

*auxins only: callus formation *cytokinins only: reduced cell division

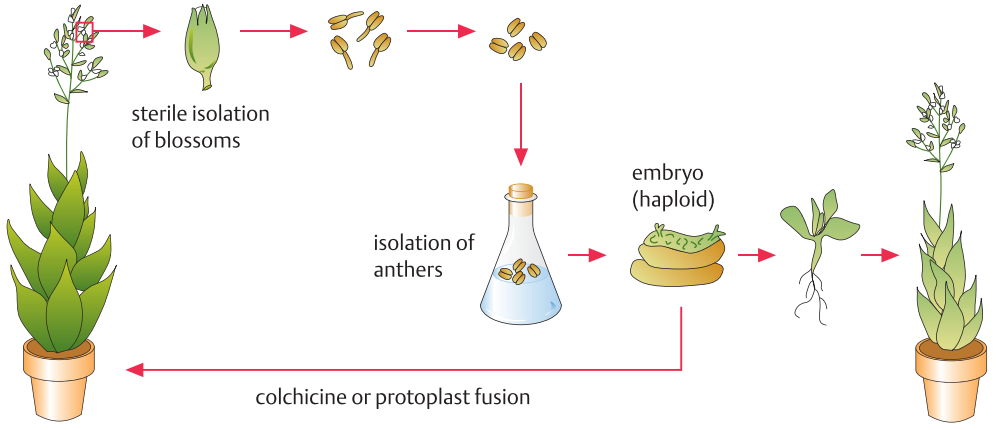
Virus-free plantlings by meristem culture

	elimination of
tomatoes	tomato aspermic virus
strawberries	nepovirus and others

	elimination of
chrysanthemum	chrysanthemum B virus, chrysanthemum stunt viroid

usually the meristem culture is heat-treated at 40 °C

Haploid culture



Plant cell suspension culture

General. Similar to microorganisms or animal cells, plant cells can be propagated under aeration in sterile liquid media to which plant hormones (→276) have been added (suspension culture). If cells from such cultures are plated on solid nutrient media, they may develop into embryoids, from which intact plants can be regenerated. Suspension cultures of plant cells are used for 1) rapid screening of variants with promising new properties; 2) preparation of protoplast cells and transgenic plants; and 3) production of secondary plant metabolites in a bioreactor.

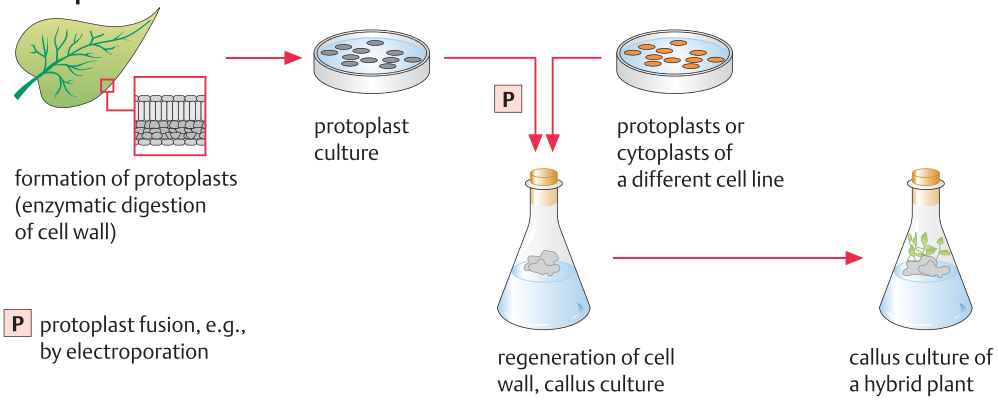
Methods. Plant cells from stock or callus cultures are transferred to a liquid medium, where they grow in a heterotrophic mode, i. e., in the presence of carbon and nitrogen sources, minerals, and plant hormones. For suspension cultures, shake flasks are used; for the production of secondary metabolites, bioreactors up to a volume of several m^3 are preferred.

Suspension cultures for screening are an excellent tool for quickly probing variants of cells for new properties, e. g., for enhanced salt tolerance, improved resistance to herbicides, or formation of secondary metabolites. Since such screening is much faster than classical selection of new variants, which is done by seeding, growth, and probing over several generations, the new method has become quite popular in plant breeding, e. g., for the improvement of soybean, citrus, sugar cane, maize, wheat, and potato varieties, but especially for the preparation of novel stress- and pathogen-tolerant plant hybrids. Some disadvantages of this method are that mutants with undesired traits may also be selected. Furthermore, the regeneration of intact plants from suspension cells is not always easy and the properties of the regenerated plants do not always correlate with the properties observed in suspension culture. Suspension cultures are widely used for screening, e. g., in the selection step of transgenic plants, which express a desired genetic trait that was introduced by targeted gene transfer.

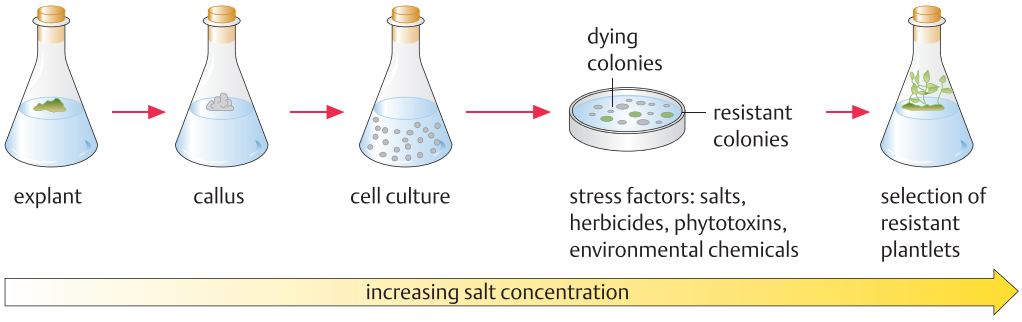
Protoplast cultures. The polysaccharide cell wall of plant cells can be removed in an isotonic solution by careful protocols using cellulases, hemicellulases, and pectinases. The resulting protoplasts can be fused with other protoplasts using chemical or electrical procedures (protoplast fusion), leading to somatic hybrids with combined genomes. If protoplasts are fused with enucleated cytoplasts, inheritable homozygous traits originating from the cytoplasmic organelles (plastids) can be transferred. This method is being successfully used for the transfer of cytoplasmic male sterility. It is highly valued in breeding, since it guarantees complete outbreeding. The fusion of protoplasts originating from different plant species has been intensely investigated as it is a method to achieve genetic recombination of plants which do not interbreed (e. g., potato + tomato = pomato), but so far this approach has not resulted in new plant varieties.

Plant cell bioreactors. The capacity of plant cells to propagate in the form of genetically omnipotent cell suspension cultures can sometimes be used for the production of valuable products. For example, secondary metabolites such as shikonin, berberin, and paclitaxel (taxol) are being industrially produced from plant suspension cultures using bioreactors up to several m^3 in volume (→94). In the case of paclitaxel, manufacturing is done with a specific cell line of *Taxus* propagated with the endophytic fungus *Penicillium raistrickii*. In a typical case of production of a secondary metabolite, phytohormone-treated cells from callus cultures are transferred to suspension cultures, increasing the reactor scale step by step. A fermenter scale of 75.000 l is being used for the production of paclitaxel from *Taxus* cell cultures. Attempts to use suspended plant cells in single-step biotransformations, e. g., in regioselective glycosidation or hydroxylation reactions, were, however not economically successful, since these reactions can often be carried out in a simpler way using enzymes or recombinant microorganisms. Stirred reactors, airlift, bubble-column, and other reactors have been investigated. Single-use bioreactors are often used for exploratory fermentation runs such as media optimization. Recovery follows standard methods used for fermentation products. Al-though the process engineering challenges involved in this technology have usually been solved, the low stability of highly productive cell lines is often a problem, aggravated by the fact that the biosynthesis of most secondary plant products (→36) and its regulation is still poorly understood.

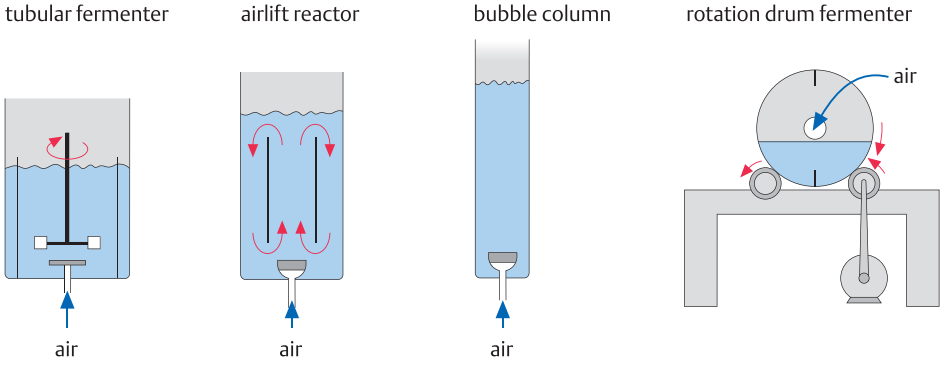
Protoplast fusion



Suspension culture for selection



Bioreactor types for plant cell cultivation



plant	product	application	scale, yields, remarks
Digitalis lanata	methyl digoxin from methyl digitoxin (12β hydroxylation)	heart-/circu-lation therapy	300 L airlift reactor, semicontinous process, ~75 % yield in 40 – 60 h
Lithospermum erythrorhizon	shikonin	cosmetics	2-stage reactor, 200 or 750 L, 23 d duration
Berberis sp.	protoberberin	pharmaceutical	up to 1.7 g/L
Panax ginseng	ginseng pieces	health	30 L bioreactor
Coleus blumei	rosemarinic acid	pharmaceutical	yields not provided
Taxus spp.	taxol	antitumor agent	up to 1 g/L after 14 d
Vanilla planifolia	vanillin	aroma compounds	16 mg/L after 45 d

Transgenic plants: methods

General. Various methods have been developed to transfer foreign DNA into a plant genome. In dicots such as tomato, tobacco, potato, pea, and bean, the Ti plasmid (→58) from *Agrobacterium tumefaciens* provides the method of choice. For dicots or monocots that can be regenerated from protoplasts, electroporation or protoplast transformation has been successfully used. Intact plant cells can be transformed by microinjection of DNA or by biolistic procedures. Successful transformation of plant cells is usually evaluated by PCR methods or by cotransformed reporter genes.

Ti plasmid. The soil bacterium *A. tumefaciens* can infect dicots, initiating cancerous cell propagation at the root neck, called root gall disease. In this process, a plasmid of ca. 200 kbp, the Ti (tumor-inducing) plasmid (→58), is transferred into the plant cells, resulting in the integration of T-DNA, a DNA fragment 15–30 kbp long, into the plant chromosomal DNA. T-DNA is flanked by imperfect tandem repeats of 25 bases (left border LB and right border RB) which are essential for their transfer. For the transformation of dicots, modified Ti plasmids have been developed which keep their infectivity but have lost their virulent properties. They contain, in addition to the gene to be transformed, the T-DNA, an origin of replication for *E. coli* or other laboratory host, and a reporter gene. Transformation is usually carried out by infecting susceptible plant cells in culture with recombinant strains of *A. tumefaciens.* Using organ-specific promoters and leader sequences enables T-DNA to be directed to any desired location in the target cell (leave, stem, root, or subcellular compartments such as chloroplasts or mitochondria). It has also been possible to induce plasmid DNA expression by external factors such as high temperature, drought, pathogen infection, light quality, or circadian cycles. A plasmid obtained from *A. rhizogenes*, the Ri plasmid, is sometimes used in a similar fashion. Plant viruses such as caulimo or gemini viruses (→6) are less well suited as vectors: they either have too low a capacity for foreign genes, too small an infection spectrum, or a replication mode that is too complex for practical use.

Transgenic monocots. Many monocots, such as wheat, barley, rice, and maize, are highly important agricultural plants, but until recently could not be transformed by *A. tumefaciens*. Recent work, however, has shown that the inducer acetosyringone can be used as a chemical signal for activation of the virulence region of the Ti plasmid in monocots. Using such protocols, cells of monocots can be transformed with the Ti plasmid system. Their regeneration into intact recombinant plants, however, poses another challenge, severely limiting all methods that depend on protoplasts, such as microinjection, electroporation, and liposome fusion. As a result, biolistics has become the method of choice. In this procedure, plant embryos are bombarded with highly accelerated tiny gold or tungsten particles on which the foreign DNA has been adsorbed. Gene transfer by this method is, however, often quite unstable ("transient expression").

Interference with gene expression. Three protocols are mainly used to obtain knockout plants. In one, integration of a foreign stretch of DNA into a target gene region through homologous recombination is used in a manner similar to the preparation of knockout animals. Unlike in animals, however, the efficiency of this method is very low in plants. Much better results are obtained if a copy of the target gene, induced by a suitable promoter, is integrated in the reverse reading direction. Transformed plants transcribe this gene into an antisense RNA, which complexes with the mRNA of the correct gene and is destroyed by RNase as a nonfunctional RNA complex (→42). Since this is a statistical process, the success rate is only in the range of 1–2 %. Genome editing is becoming a 3d pillar of plant transformation (→64).

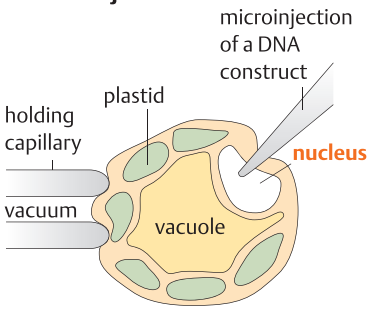
Plant genomes. The first plant genome to be completely sequenced was *Arabidopsis thaliana*, with a genome size of ca. 100 Mbp, distributed among 5 chromosomes. More recently, genome sequencing of two varieties of rice *(Oryza sativa)*, with ca. 420 Mbp on 12 chromosomes, and of several other plant genomes such as soya, potato, wheat, rye, cotton, tobacco, strawberries etc. was completed. Difficulties in analyzing plant genomes include their size (maize: 2.5 Gbp), their polyploidy (wheat: hexaploid; strawberry: octoploid), and alignment of their repetitive sequences (wheat: hexaploid) and transposons.

Transformation methods

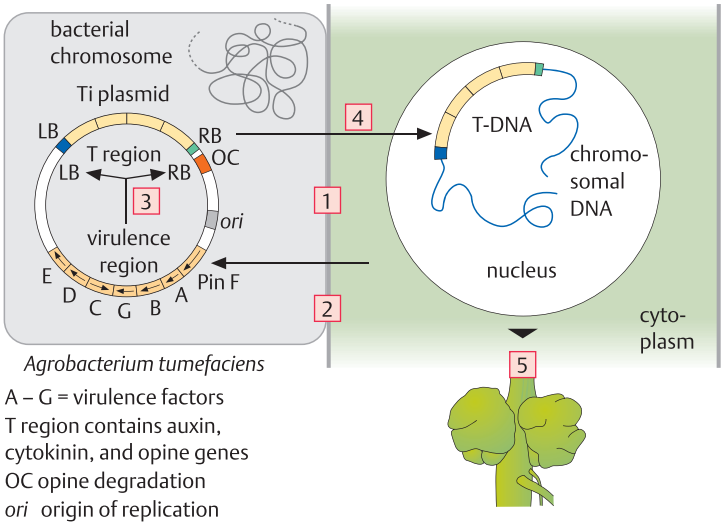
	monocots	dicots
microinjection (protoplasts)	+	+
transfection (protoplasts, Ca^{2+})	0	+
electroporation (protoplasts)	0	+
biolistics	+ +	+
T-DNA	+	++
plant viruses	0	0

0 occasionally successful + possible ++ method of choice

Microinjection



Infection of a plant with *Agrobacterium tumefaciens*



Agrobacterium tumefaciens

A – G = virulence factors
T region contains auxin, cytokinin, and opine genes
OC opine degradation
ori origin of replication

1 *A. tumefaciens* attaches to an injured plant cell

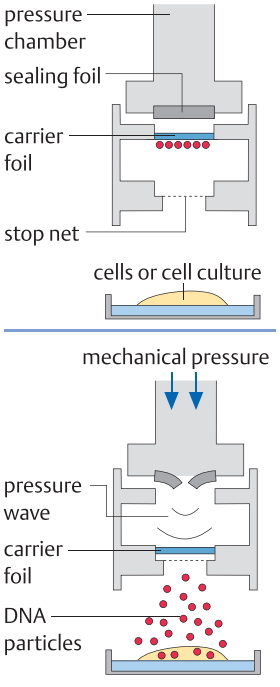
2 plant cell sends chemotactic signals and activates virulence region

3 virulence region of Ti plasmid activates T region

4 T region is integrated into plant chromosome as single-stranded DNA

5 auxins and cytokinins initiate tumor growth. Opine becomes C-source for *A. tumefaciens*

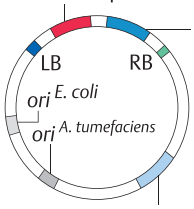
Biolistics



Cloning vectors derived from *A. tumefaciens*

binary vector

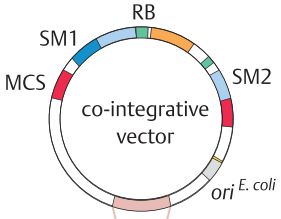
MCS = multiple cloning site



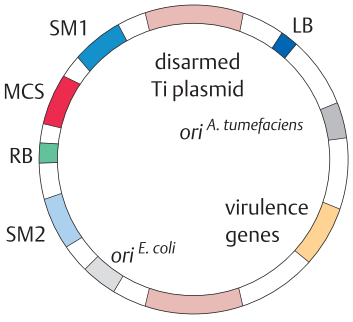
selection marker for plant (SM1)

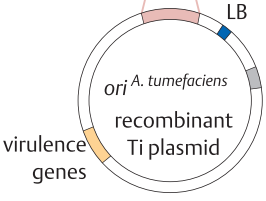
selection marker (SM2) for *E. coli* and *A. tumefaciens*

cointegrative vector system



recombination of homologous DNA-sequences





Transgenic plants: resistance

General. In the USA, > 30 transgenic plants have been registered for agricultural use and are being raised on 73 million ha of land. This includes transgenic cotton, potato, maize, rapeseed, soybean, and tomato. In most cases, foreign genes that convey tolerance to herbicides, insecticides, or viruses were incorporated. On a global scale, in 2014 >10 transgenic crops were cultured in 28 nations on 181.5 million ha of land. Plants with enhanced stress tolerance or ornamental plants with modified colors are in an advanced stage of development.

Herbicide-tolerant plants. About 10 % of all harvests are lost due to weeds. The ideal herbicide should be active at low concentrations, not inhibit the growth of the agricultural plant, be readily degraded, and not reach groundwater. Transgenic plants reduce the need for herbicides, since they are made genetically resistant by 1) containing the herbicide-sensitive protein in excess, 2) showing reduced binding of the herbicide, or 3) inactivating the herbicide through biodegradation. As an example, soybeans resistant to the broad-spectrum herbicide glyphosate (Roundup™) were produced by isolating glyphosate-resistant strains of *E. coli*, isolating the bacterial gene coding for 5-O-enolpyruvyl shikimic acid-3-phosphate synthase (ESPS synthase, the herbicide target), and expressing it in soybean under the control of a plant promoter. Resistance of tobacco, potato, rapeseed, and other plants to phosphinothricin (→210) (Basta™), an inhibitor of glutamine synthetase, was achieved by expressing a phosphinothricin acetyltransferase (PAT) from *Streptomyces hygroscopicus*.

Insect-resistant plants. *Bacillus thuringensis* synthesizes a protein of M_R 250 kDa (δ-endotoxin, BT toxin), which forms a highly toxic protein in insect intestines after proteolysis. In plants and mammals, this transformation does not occur. Thus, BT protein has been expressed in numerous plants as an insecticidal protein. Using optimized codons and strong constitutive promoters, e. g., for the 35S protein of cauliflower mosaic virus, the expression rate was increased ca. 1000 fold. Cloned protease inhibitors were also successfully used as insect control agents.

Fungus-resistant plants. Fungal infections lead to important damages to crops. An important historical example is the potato blight (caused by *Phytophthora infestans*), which in the 19th century led to famines throughout Europe, especially in Ireland. By overexpressing glucanases or chitinases of plant origin, directed against the fungal cell wall, the fungal resistance of tobacco was increased. Good success was also obtained with ribosome-inactivated proteins (RIP).

Virus-resistant plants. Viruses (→6) can also lead to significant harvest losses, e. g., potato virus 4 in potato or *Rhizomania* virus in sugar beet. Attempts are being made to interfere with virus replication by expressing nonfunctional virus capsid proteins in plants (cross protection). The expression of antiviral antibodies or of hammerhead ribozymes is also being investigated. (→42)

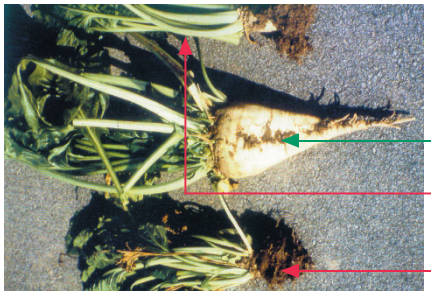
Plants with enhanced stress tolerance. Many forms of physiological stress (strong light, UV radiation, heat, drought) are accompanied by the formation of oxygen radicals, in particular of the oxygen radical anion. Transformed plants, expressing the gene product superoxide dismutase under control of the 35S promoter of cauliflower mosaic virus, not only were more resistant to physiological stress, but also wilted more slowly.

Altered blossom colors, aging. Form and color are important characteristics of ornamental plants, as storage stability and aroma are of fruits. Incorporating foreign genes or silencing indigenous genes can affect these properties. For example, genes involved in the secondary metabolism of other plants have been expressed to modulate chromophore biosynthesis, mostly within flavonoid or anthocyan glycoside metabolism. In the "blue" rose, a P450 monooxygenase from pansy, which hydroxylates dihydroquercetin yielding delphinidin, a blue pigment, is expressed, and the gene coding for dihydroflavonol-4-reductase, committed to the biosynthesis of (red) anthocyanins, was silenced. The deep-blue pigment of cornflower, protocyanin, is a complex of six pigments bound by Fe^{2+}, Mg^{2+} and Ca^{2+} and could not yet be mimicked in the "blue" rose which, in fact, has a violet color. To silence genes, the antisense technique has been successfully applied: another commercial example is the Flavr-Savr™ tomato in which a pectinase gene has been inactivated, leading to a longer shelf life. However, this project was abandoned.

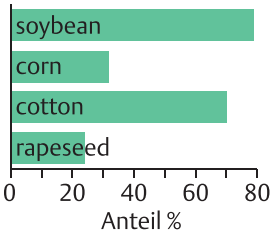
Resistant plants (examples)

plant	overexpressed/ foreign/modified or inactivated gene product	desired property	trans-formation method	company organization
cotton	acetolactate synthase	resistance to sulfonylurea		DuPont
maize	BT-δ-endotoxin	resistance to European corn borer		DeKalb, Monsanto
	glutamine synthase, transacetylase, EPSP synthase	phosphinothricine and glyphosate resistance		Bayer Crop Science, Zeneca Seeds, Monsanto
soy-bean	glutamine synthase, transacetylase, EPSP synthase	phosphinothricine and glyphosate resistance		Bayer Crop Science, Zeneca Seeds Monsanto
potato		resistance to Colorado beetle		Monsanto
	polyphenol oxidase	prevention of browning	antisense construct	Bayer Crop Science, Zeneca Seeds, Monsanto
papaya		papaya ringspot virus resistance		Cornell-University/ University of Hawaii
tomato	polygalacturonase	retarded maturation	antisense construct	Calgene, Monsanto "Flavr Savr"
		thicker skin, retarded maturation	antisense construct	Zeneca Seeds
	monellin	enhanced sweetness	Ti plasmid	
petunia	dihydrokaempferol reductase from maize	ruffled petals	sense construct	Max-Planck-Institute
rose	dihydroquercetin-5'-hydroxylase	blue pigment	sense construct	Suntory, Calgene

Rhizomania of the sugar beet



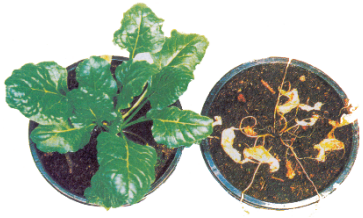
Cultivated transgenic plants



2014: 181,5 Mio. ha in 28 nations, including 73 Mill. ha in the USA

~ 50 % of all transgenic plants carry herbicide resistance

Phosphinothricin-resistant soybean



By expressing phosphinothricine acetyltransferase from *Streptomyces hygroscopicus*, the plant becomes resistant to the herbicide phosphinothricine (Basta™).
Right panel: a non-resistant control.

The blue rose

Transgenic plants: products

General. Modification of plants by genetic engineering includes not only the improvement of resistance against herbicides and pathogens (→282), but also the modification of plant components, which can be used as renewable resources, and even the synthesis of valuable chemicals in transgenic plants. Examples include 1) modification of the amino acid, starch, lignin, or oil composition; 2) expression of antigens, antigen fragments (plantibodies), vaccines, human serum albumen, or biopolymers.

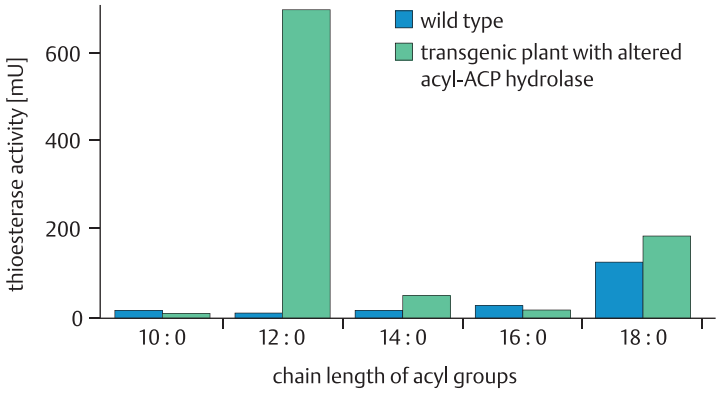
Modification of plant chemicals. Several routes have been explored for supplementing plant proteins with essential amino acids (→124) required in human food (usually L-lysine and L-methionine): 1) expression of proteins from other plants, which have a more adequate composition; 2) site-directed mutagenesis of indigenous storage proteins, replacing non-essential with essential amino acids; and 3) cloning genes for deregulated key enzymes in branched metabolic pathways, e. g., aspartokinase from *Escherichia coli* and dehydrodipicolinic acid synthase from *Corynebacterium* for enhanced synthesis of L-lysine (→128). ADP glucose pyrophosphorylase, a key enzyme of starch biosynthesis, is an allosteric enzyme in most plants. To increase the percentage and modulate the composition of starch, a nonregulated enzyme from *E. coli* was expressed in tomatoes, resulting in fruits with a 20 % higher starch content. The ratio of linear amylose to branched amylopectin is essential to the properties of starch in processed food and in technical applications. Expression of a gene that is responsible for α-1,6 bond formation, *glgB* from *E. coli*, in potatoes and under control of the granule-bound starch synthase promoter (GBSS promoter), led to the formation of starch with a 25 % higher percentage of amylopectin. In the Amflora™ potato developed by BASF, the formation of amylase is completely stopped, because the transcription of GBSS has been halted using a specific antisense-RNA (→42, 46). Starch from this potato is instead made up from amylopektin, which is an industrial raw material, e. g. for wallpaper paste. The fatty acid composition of oil plants was also significantly modified. Thus, lauric acid (C12), an important renewable raw material for the manufacture of surfactants soluble in cold water, occurs only in tropical oils (palm kernel oil, coconut oil) as a triglyceride. By cloning chain-length-specific fatty acid ACP thioesterases from laurel (*Umbellularia californica*) into suitable rapeseed lines, complemented by other measures, transgenic rapeseed varieties were obtained whose seed contains 50 mole % of trilauroyl glycerol (→162). The value of the world's total annual timber harvest is in excess of 400 billion US$. In view of faster growing trees and reduced pollution during the paper making process, the lignin content of woods was reduced by modulating genes required for lignin biosynthesis, e. g., in the cinnamic acid pathway (Coumaryl-3'-Hydroxylase). Fast growing hardwood trees like poplars and eucalyptus trees are the focus of such investigations (→184).

Expression of valuable chemicals. Most of these experiments were carried out with transgenic tobacco plants or in *Arabidopsis thaliana*, since transformation of these organisms is relatively easy. The expression of human serum albumen was achieved with good yield, as was the expression of complete IgG antibodies ("plantibodies") (→244), which, in view of developing an immunological means of preventing tooth decay, were directed against the adhesin of *Streptococcus mutans*. Concentrations of up to 1 % of total protein were obtained. The economic expression of antigens in plants has been discussed for carrying out inexpensive vaccination procedures in developing countries, using food intake as the vaccination step. In model experiments, a fragment of the surface antigen of hepatitis B virus (→250) was expressed in tobacco at levels of ca. 0.01 % of the total soluble protein, and feeding mice with meal derived from such tobacco plants led to an immune response. Eating potatoes containing expressed fragments of a heat-labile enterotoxin B from *E. coli*, a protein leading to diarrhea, also led to an immune response in human volunteers. When a three gene *Ralstonia eutropha* operon coding for the synthesis of polyhydroxybutyric acid (→154) is expressed in the chloroplasts of *Arabidopsis thaliana* or rapeseed, this valuable polymer can be produced by photosynthesis. The economic processing of plants containing this polymer must, however, be further optimized, and a critical European public still needs to be convinced.

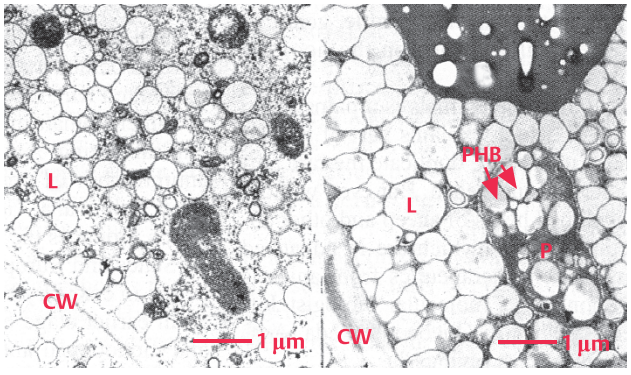
Tomato puree

from tomatoes with enhanced sugar content

Transgenic rapeseed with altered fatty acid composition



Polyhydroxybutyric acid in the plastids of transgenic rapeseed



CW = cell wall
P = plastid
PHB = polyhydroxybutyric acid
L = lipid

Plants as bioreactors

plant	foreign/modified/inactivated gene	observed property	transformation method
modification of indigenous plant genes			
rapeseed, soybean	derepressed aspartokinase (*E. coli*) and dihydrodipicolonic acid synthase (Corynebacterium)	higher lysine content	Ti plasmid
potato, tomato	starch synthase	altered starch composition	antisense construct
rapeseed	acyl-ACP hydrolase	altered fatty acid profile	Ti plasmid
Arabidopsis	γ-tocopherol methylase	synthesis of vitamin E	Ti plasmid
pine tree	shikimic acid synthase	altered lignin content	antisense construct
expression of genes foreign to plants			**yields [g/kg]**
potato	coat protein of hepatitis B virus	immune response in mice	
	human serum albumen	formation of human serum albumen	
tobacco	IgG fragments against adhesin of *Streptococcus mutans*	formation of antibody fragments	10
Arabidopsis rapeseed	3 genes (*phb* operon) from *Ralstonia eutropha*	formation of polyhydroxy butyric or valeric acid	140

transgenic plant (operon for PHV) → **harvest, extraction** → **recovery** → up to 140 g/kg "phytopolymer"

Aerobic wastewater treatment

General. Aerobic wastewater treatment was introduced about 100 years ago, using trickling filters and aeration basins. In combination with the construction of sewage systems, which were started in ca. 1850, these measures resulted in a dramatic increase in human life expectancy, due to a decrease in epidemics. Today, in most industrial countries almost all wastewater is biologically treated before entering surface waters: in Germany (2007), for example, ca. 10,000 municipal sewage plants perform biological treatment. However, worldwide there remains much to do, e. g., in many coastal regions, where considerable amounts of unpurified wastewater are still entering the sea, and in those developing countries where industrialization and rapid population growth coincide.

Composition of wastewater. Domestic and industrial wastewater have different compositions. Although the latter varies depending on the type of industry producing it, domestic wastewater is surprisingly constant in composition, after temporary deviations have been averaged. Taking German standards as an example, it contains an organic load of 60 g BOD_5 per inhabitant per day ("inhabitant equivalent"). BOD_5 (biochemical oxygen demand after 5 days) is measured as oxygen consumption using a standardized procedure. Other important parameters are the COD (chemical oxygen demand) and the TOC (total organic carbon), which are determined by chemical oxidation of a wastewater sample and thus represent the total organic load as well as the oxidizable inorganic components of a wastewater sample. Further important parameters are the total and the organic nitrogen load, and the phosphate load. Many wastewaters are of mixed domestic and industrial origin.

Microbiological aspects. In the sludge of an aerobic wastewater treatment plant, a large variety of microorganisms, algae, and protozoa form a symbiotic microcosm (biocenosis) (→74). If only domestic wastewater is treated, these populations are quite constant over long periods of time. The addition of industrial wastewater may dramatically alter the composition of these populations. Recently, sludge starter cultures (→114) have become available which are specialized for the oxidation of certain industrial waste components and thus aid the purification of industrial wastewater.

Trickling filter process. In this process, mechanically pretreated wastewater is percolated through a tower filled with materials having large surface areas (e. g., lava stones). As a result, filamentous and polysaccharide-secreting bacteria form sludge consortia which aggregate on the surface of these materials, oxidizing wastewater components in a continuous manner. The oxidative capacity of the system is limited by the diffusion of oxygen. Excess sludge is removed by flushing under pressure and further treated by anaerobic sludge digestion.

Aeration basins (aeration tanks). In this process, wastewater is retained in a stirred aerated basin. This technology is often preferred over trickling filters because its oxidation efficiency is ca. 5 times higher, due to the active introduction of air into the process. Performance can be further increased by injecting oxygen instead of air. The sludge formed in this process, composed of similar consortia as on trickling filters (e. g., *Achromobacter, Aerobacter, Alcaligenes, Bacillus, Citromonas, Escherichia, Pseudomonas, Zoogloaea*) is removed in a subsequent sedimentation basin and transferred to anaerobic sludge digestion. Disadvantages of this procedure are its open construction, resulting in limitations for process modifications and in odor nuisances for the neighborhood. In some countries such as Germany, tertiary treatment of sewage has become standard. It is used to remove phosphate, a eutrophication factor, by precipitation or biological processes, and nitrate, a risk factor in drinking water, by a biological process using consortia of nitrifying and nitrate-reducing (denitrifying) bacteria.

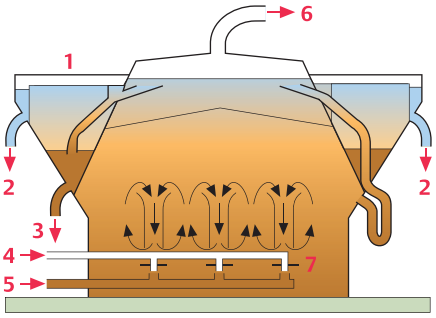
Tower biology. If large volumes of industrial wastewater must be purified, a highly engineered process involving closed fermenter towers can be used. In technologies developed by Hoechst and Bayer in Germany, 30-m high towers with loop pumps are used (→196). Aeration is enhanced by thorough optimization of the sewage feed and air impellers. The oxidative capacity of such towers is ca. 50-fold higher than that of aeration basins, and odors are retained inside the closed system. Domestic wastewater may be added to enhance overall biodegradation.

Composition of waste water

origin	inhabitant equivalent*	remarks
domestic wastewater	1	per inhabitant
breweries	150 – 350	per 1,000 L beer
dairy (without cheese production)	25 – 70	per 1,000 L milk
starch factory	500 – 900	per 1,000 t corn wool
wool	200 – 4 500	per t wool
paper factory	200 – 900	per t paper
sugar factory	1,000 – 2,000	per t sugar

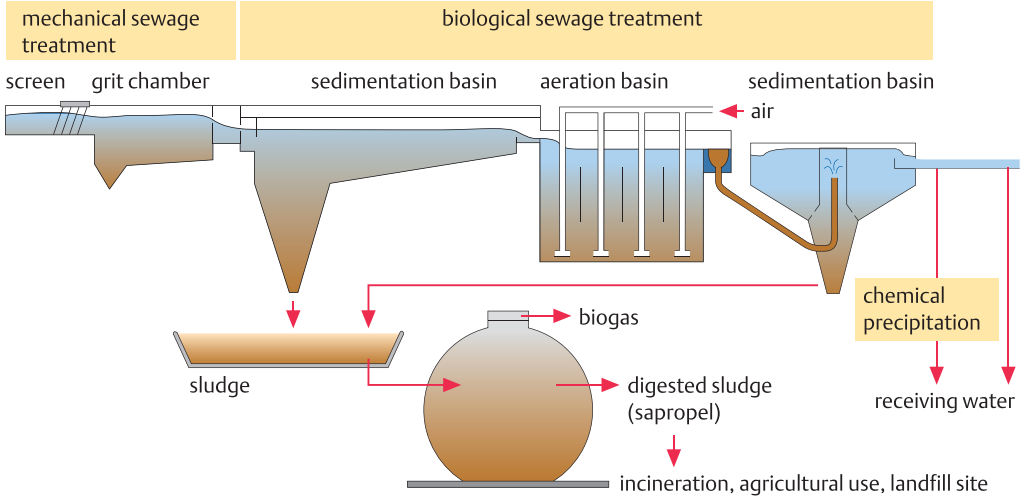
*one inhabitant equivalent corresponds to 60 g BOD_5

Tower biology*

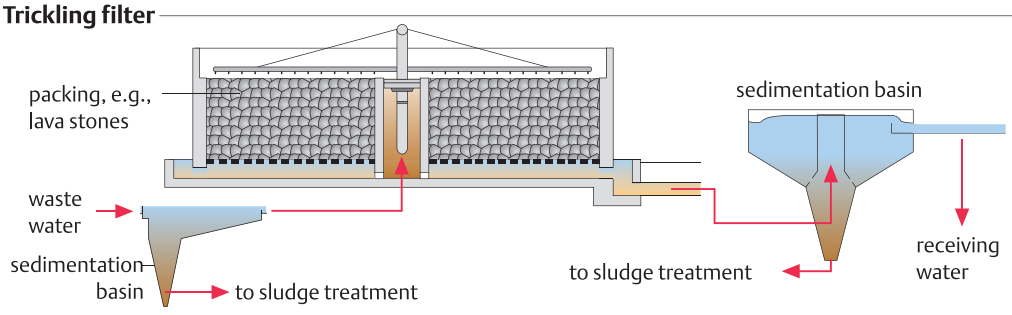


1 sedimentation tank
2 purified wastewater
3 excess sludge
4 compressed air
5 wastewater
6 exhaust air
7 injector
***e.g. Bayer AG**

Biological sewage plant



Trickling filter



Typical construction and performance data

	trickling filter	aeration basin	tower biology
height or depth (m)	2 – 4	3 – 6	30
diameter (m)	up to 30	up to 30	30
working volume (m^3)	~ 10	~ 100	15,000
residence time (h)	~ 4	6 –10	14
reduction in BOD_5 (t/d)	0.5	2	100

Anaerobic wastewater and sludge treatment

General. Sludges formed during aerobic wastewater treatment are usually subjected to anaerobic fouling before they are incinerated, deposited in landfills, or used in agriculture. Anaerobic sludge digestion is probably the highest-volume process in biotechnology. In Germany alone, ca. 5,000 sludge treatment units with a total volume of 1 million m^3 produce 100 million m^3 biogas per year. Wastewater can be treated anaerobically as well, using anaerobic fluid bed reactors (→86). In some developing countries such as China and India, the removal of waste by anaerobic digestion in small decentralized facilities has been implemented to produce biogas (methane) as a local energy source (→330).

Microbiological aspects. Methane is formed from sludge by the action of three types of bacterial consortia (→10): 1) various obligatory or facultative anaerobic bacteria (Clostridia, Enterobacteria, Streptococci) degrade starch, fats, and proteins to organic acids, H_2, and CO_2; 2) acetogenic bacteria transform higher fatty acids into acetic acid, H_2, and CO_2; and 3) methanogenic bacteria form methane and CO_2 from acetic acid; the latter are obligate anaerobic strains and frequently belong to the Archaebacteria.

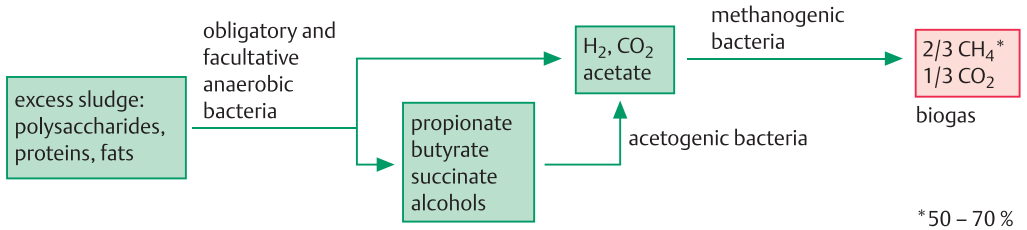
Analytical aspects. The essential parameters in sludge digestion are the reduction of organic carbon (TOC, DOC, or COD), and biogas formation. Total organic carbon (TOC) is determined by infrared spectroscopy after persulfate oxidation of a sample, dissolved organic carbon (DOC) as TOC after filtration, and chemical oxygen demand (COD) by oxidation with dichromate, followed by titration. All three methods thus represent the amount of CO_2 formed or O_2 consumed that is required for the complete chemical oxidation of a sample (here, sludge), or its dissolved fraction (DOC), to CO_2 and H_2O. Biogas is the end product of anaerobic sludge digestion. It is composed of ca. ⅔ CH_4 and ca. ⅓ CO_2, with traces of H_2, N_2, H_2S, and other gases. Its composition is usually determined by gas chromatography.

Technical aspects. Compared with the aerobic treatment of wastewater, anaerobic digestion of sludge is considerably slower (residence time ca. 20 d). The development of the microbial consortia required for sludge digestion occurs only in a very narrow range of pH and temperature, necessitating thorough control of these parameters. The advantage of this process is, however, that the sludge is nearly completely (> 90 %) transformed into CH_4 and CO_2 with the formation of little biomass and with little odor. The digested sludge remaining after fouling is incinerated, deposited in landfills, or used as an agricultural fertilizer. The energy content of the biogas produced largely exceeds the energy requirements of the whole sewage-treatment plant.

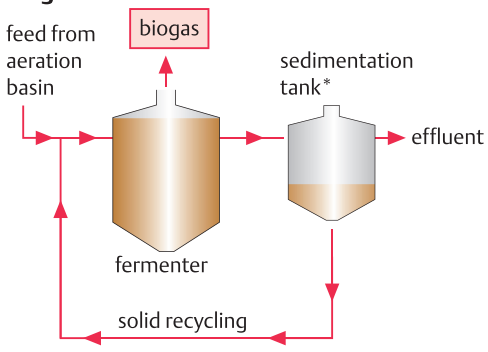
Anaerobic fluid bed reactor. Wastewater with a high load of biodegradable organic residues can be advantageously treated by anaerobic sludge treatment, i. e., without prior aerobic oxidation. In such processes, tower reactors (→96) are used where the microbial consortia aggregate into sludge particles, either on their own or after addition of particles with a high surface area. These particles have a high density and thus tend to sediment, resulting in the enrichment of bacteria in the lower part of the tower. Since the wastewater is added from the bottom, and the biogas formed leads to further mixing of the column, the process operates with high efficiency. In the upper part of the column, gas is separated from the particles by a gas separator. With some wastewaters, e. g., from paper mills or the sugar and starch industries, a TOC reduction > 95 % and excellent biogas production can be achieved within less than a day even when loads are high.

Biogas. In China alone, more than 7.6 million households run biogas digesters, which can generate 200 million $m^3\ y^{-1}$ of biogas and provide enough fertilizer to reduce the consumption of firewood significantly. Biogas plants are based on simple technology. Sludges are mostly derived from agricultural biomass such as manure, harvest residues, or domestic waste. In the context of a future turnaraound in energy generation from renewable resources (bioeconomy), biogas production has found renewed interest in industrial countries, and this technology is being further developed. As of 2013, some 7,700 biogas plants were installed in Germany alone, producing 2.2 billion kWh of electricity and feeding 500 million m^3 of biomethane into the German natural gas network.

Degradation reactions

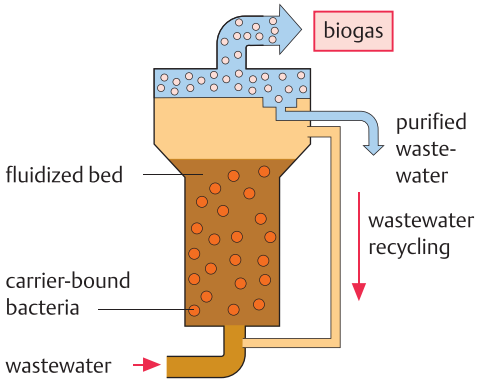


Digestor



*frequently incorporated in digestor

Anaerobic fluidized bed reactor



Typical construction and performance data

	conventional digestor	contact fouling	fluidized bed reactor
height, volume	up to 30 m, 16 000 m^3	pilot plant stage	up to 20 m 2 000 m^3
load (kg COD/m^3 · d)	1 – 8	1 – 5	5 – 30
residence time of liquid (d)	10 – 30	0.5 – 25	0.2 – 1.5
residence time of microorganisms (h)	10 – 30	> 20	> 100
reduction in COD (%)	30 – 70	60 – 90	80 – 90

Biogas plants and agriculture

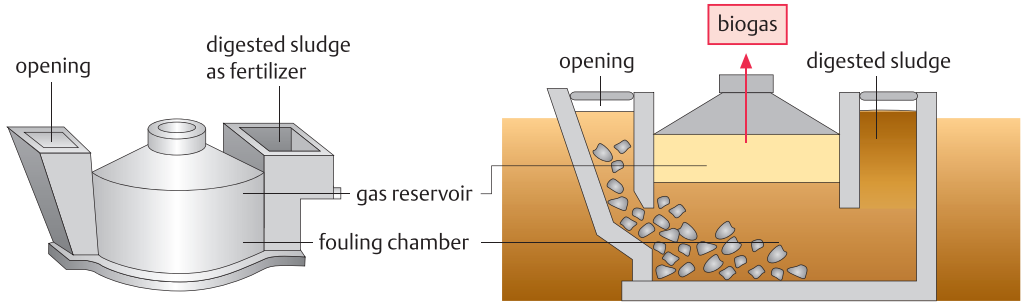
biomass
excrements, liquid manure, harvest residues, waste materials

sedimentation tank

fouling chamber
C/N ~ 30:1; 20 – 60 d, 30 – 35 °C

60 – 70 % conversion to biogas*

*2 – 5 m^3 biogas/d from 1000 kg domestic animal, ~ equivalent to 1 – 3 L diesel oil



Biological treatment of exhaust air

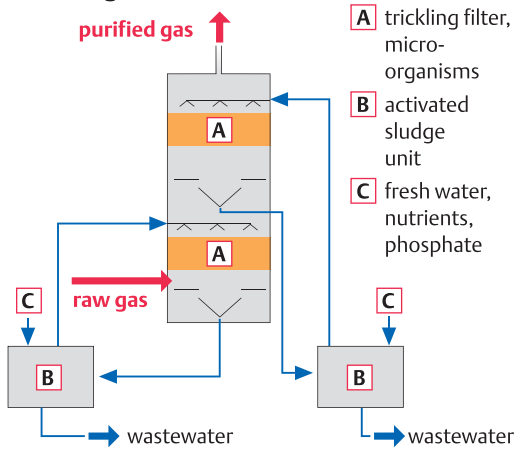
General. With ever-intensifying legislation concerning pollution control and waste gas emissions, new ways to remove volatile organic compounds (VOCs), such as those causing odors, from exhaust air have been explored. Microorganisms are able to degrade such components, after they have been absorbed in an aqueous phase. They are used in biowashers to oxidize water-soluble compounds and in biofilters to remove sparingly water-soluble organic waste gas components.

Exhaust air and gas. Biofilters have been used for many decades to remove smells from some parts of sewage plants (→286), in particular from sludge thickeners or dewatering units of anaerobic digesters. Exhaust air from many other industrial and agricultural plants such as foundries, rubber and polymer-producing plants, varnish firms, breweries, food factories, chicken and pig farms, and slaughterhouses is also being successfully treated. Typical organic odor components are the lower fatty acids, amines, and mercaptans (manure, animal rendering plants); phenols and low-molecular-weight amines; aldehydes and ketones (foundries); aromatic compounds (varnish firms); and furfural (food factories). Biofilters are sometimes also used in the cleanup of soils at disposal sites and landfills with a low load of organics. For large volumes of malodorous air, a biofilter may be the only cost-effective solution.

Biofilters are usually units of rather simple construction. Compost, barks, peat, or wood with a high surface to volume are being used as carriers. Inorganic materials such as cinders or lava are added to reduce the packing density. If humid exhaust gas is passed through such filters, consortia of microorganisms start to grow and to oxidize odor components in the organic exhaust gas. After several years of use, the organic support materials of the biofilter may become biologically oxidized as well, resulting in reduced performance and increased pressure drop. With a new packing, the previous level of performance can be recovered. Depending on their physical form, single- and multi-level biofilters can be distinguished.

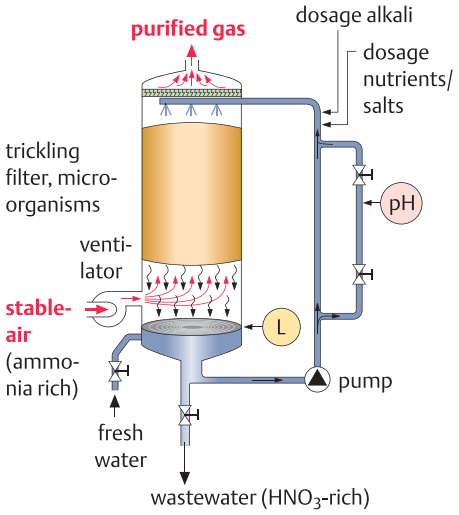
Bioscrubbers have a more complex construction than biofilters: the organic odor components are first absorbed in an aqueous phase, often enriched with nutrients, to be later metabolized in a trickling filter, in an aeration basin, or a bioreactor by adapted microorganisms. More sophisticated monitoring and control of plant performance is required in such two-step setups, which also leads to higher performance than in simple biofilters, since toxic metabolites or end products are continuously removed. If exhaust air containing H_2S is purified, rapid acidification takes place through the enrichment of *Thiobacilli* in the microbial consortium that oxidize H_2S to H_2SO_4. Bioscrubbers that are equipped with a pH control allow the generated acid to be neutralized and thus have a long operational lifetime. Their energy consumption remains quite low due to the limited amount of water consumed. For easily biodegradable solvents such as the lower alcohols, the residence time of the exhaust gas in the washer can be limited to 1–2 minutes. If a mixture of easily and poorly degradable components has to be treated, 2-stage scrubbers or a combination of a trickling filter with a washer are often used. For example, the exhaust gas from a lacquer plant can be treated in a 2-stage process where, in the first stage, the easily degradable alcohol and esters, and in the second stage the less water-soluble and biodegradable aromatic components such as toluene or xylene are metabolized. The recycling of exhaust air from buildings used in intense animal production not only serves to remove CO_2 and to replenish O_2, but also to regulate temperature, remove germs, and eliminate strong-smelling exhaust gas components, in particular, ammonia. Removal of ammonia takes place in a biowasher in which nitrifying bacteria oxidize ammonia to nitrate. The nitric acid is absorbed in water and neutralized, and the heat generated in this process is recovered with a heat exchanger. Typically, the ammonia concentration of exhaust air can be reduced to 2–4 ppm, and the emission of ammonia into the neighborhood can be reduced to 0.2 kg per animal per year, as compared to 5.3–5.6 kg from a farm not using biowashers. Modern plants are monitored and operated by measuring and control devices such as humidity sensors in the filter material.

Two-stage bioscrubber



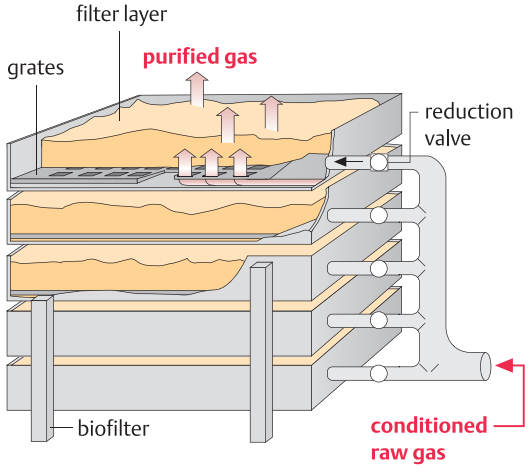
two-stage trickling filter for a meat-rendering plant: odor reduction ~ 99 %

Bioscrubber for farm emissions



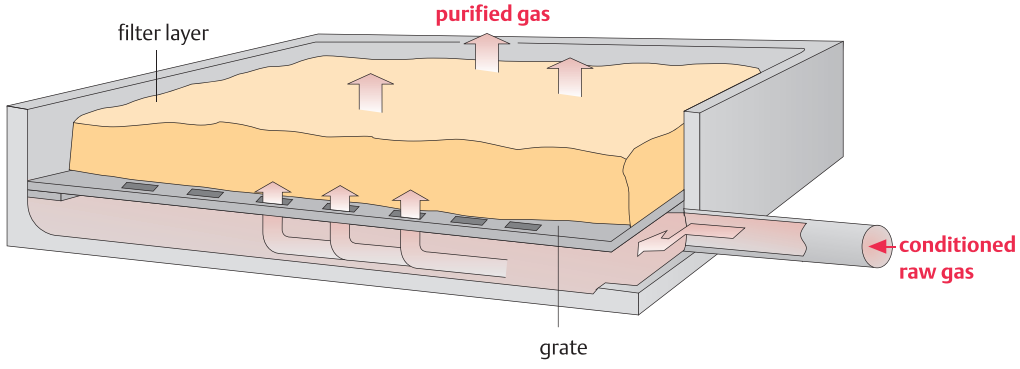
Biofilter

multiple level biofilter



filter layer:
compost, bark, peat, lava, cinders

single level biofilter



Typical components of exhaust air

source	main components	cleaning procedure
animal farm	lower fatty acids, ammonia	biofilters
industrial emissions	phenol	biofilters with Pseudo-monads
foundries	phenol, form-aldehyde, amines, ketones	biowashers
animal-rendering plants	lower fatty acids	biowashers
chemical plants	toluene, ammonia, aldehydes	biofilters or biowashers
digestors	H_2S	biowashers

Biological soil treatment

General. Consortia of microorganisms play a crucial role in establishing an ecological balance in the environment through the degradation of biomass and the mineralization of organic matter. Although this capacity has been used for a century in wastewater treatment, the microbial decontamination of soil (bioremediation) has been investigated only in the past ca. 30 years. Processes comprise the injection of microbial cultures *in situ* and the cleaning of soils after excavation. They compete with chemical and thermal processes. For biological soil treatment, natural or recombinant microorganisms can be used, but the uncontrolled release of engineered microorganism into the environment is presently not authorized (→334).

Contamination and soil structure. Anthropogenic contaminants are mainly grouped into: 1) mineral hydrocarbons (MHC); 2) benzene, toluene, xylene, and ethyl benzene (BTXE); 3) polyaromatic hydrocarbons (PAH); 4) chlorinated hydrocarbons (CHC); and, 5) trinitrotoluene (TNT), in military areas. MHC and BTXE are usually easily biodegradable. Higher condensed PAH and CHC, in contrast, are not easily biodegraded. In any assessment of the biodegradability of such compounds, the soil composition must be considered. For example, sandy soils which can be easily penetrated are easier to purify than clay soils. So far, TNT can only be immobilized by applying a sequence of anaerobic and aerobic steps.

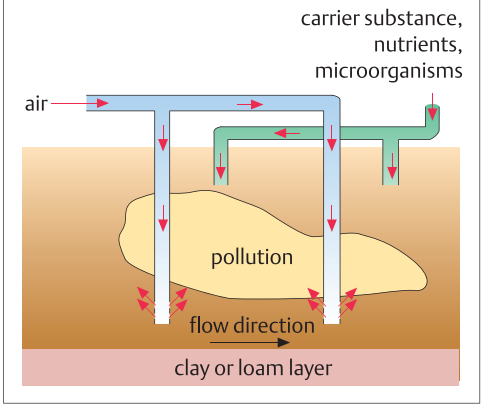
Soil treatment in situ. In this type of technology, nutrients and microbial cultures of high biodegradation activity, which have been isolated after enrichment in the presence of the contaminants, are added in solution through injection wells. Aeration is passive and ensured by, e. g., ground water mobility. The method works well for polluted soil layers up to 1.5 m thick and can be further improved by active aeration through drainage systems or by pressurized aeration. If some of the pollutants are volatile, waste gas must be carefully monitored, and suction aeration may be the method of choice. If the soil is mainly composed of clay and loam, oxygenation is not possible. In such cases, nitrate has been used as an electron acceptor. The ecobalance of such processes, however, is doubtful in view of groundwater pollution with nitrate.

Soil treatment ex situ. Soils that are contaminated with readily biodegradable chemicals are usually treated after excavation in prepared long bed reactors. Suitable microorganisms are selected in precultures for high biodegradation efficiency and later used to inoculate the soil. Nutrients are added. Upon good aeration and mixing of the piles, which are ca. 2 m high, contaminants can be reduced by > 90 % within 2 weeks. Treatment costs amount to 75–150 € m^{-3} of soil.

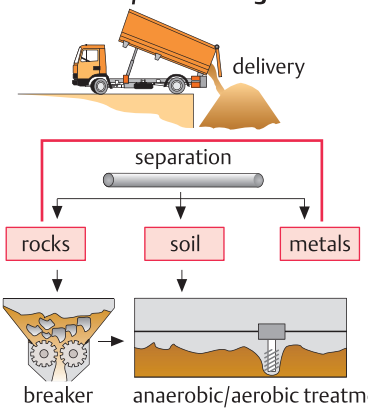
Humification of TNT. Xenobiotics such as TNT, trichloro- or tetrachloroethylene, being substituted with a considerable number of electronegative substituents, are quite difficult to metabolize under aerobic conditions. They are rather well metabolized, however, by anaerobic bacteria. As a result, a procedure was developed for degradation of the TNT-contaminated soils of military camps which is based on a combination of both methods. In the first step, a reactor filled with 25 t of TNT-containing soil was kept anaerobic by adding saccharose as an electron donor. After 18 d, TNT was reduced to triaminotoluene to such an extent that, by aeration, it is covalently and irreversibly bound to soil components such as humic acids.

Recombinant microorganisms. Metabolic steps from different microorganisms can be combined by genetic engineering techniques to provide recombinant microorganisms that are better suited for the degradation of recalcitrant chemicals than the wild-type strains. This method has been explored with considerable success for the degradation of chlorinated aromatic and aliphatic compounds such as chlorobenzoic acids, chlorobenzofurans, and aliphatic CHCs. Pseudomonas strains (→20) are often used, since they contribute to the biodegradation of aliphatic and aromatic hydrocarbons in the environment to a significant extent. They often carry part of the genetic information for these steps in the form of plasmids. An example is the toluene-degrading plasmid TOL (→58) first found in *Pseudomonas putida*. Experiments in contained bed reactors were promising, but the risks of releasing genetically modified organisms (GEMs) into the environment are still under public debate (→336).

Biological soil remediation *in situ*



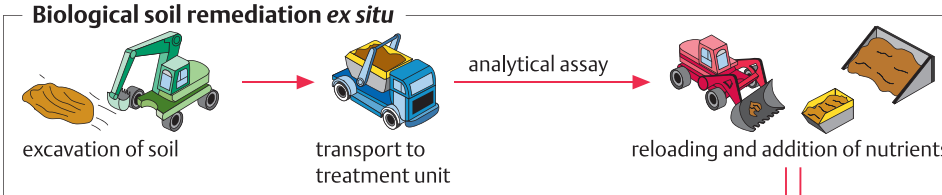
Anaerobic/aerobic degradation of TNT



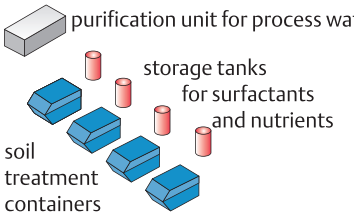
anaerobic: nutrients, sulfate, 30 – 60 d TNT → tri-aminotoluene

aerobic: nutrients, O_2, 15 d, triamino-toluene → humification

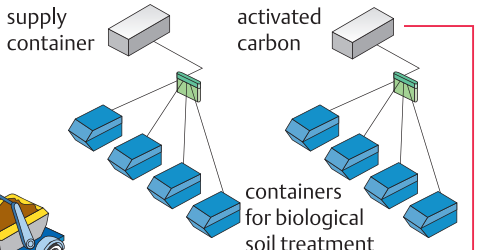
Biological soil remediation *ex situ*

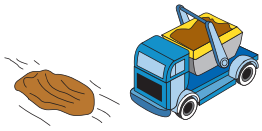


chemical-physical soil remediation



biological soil remediation

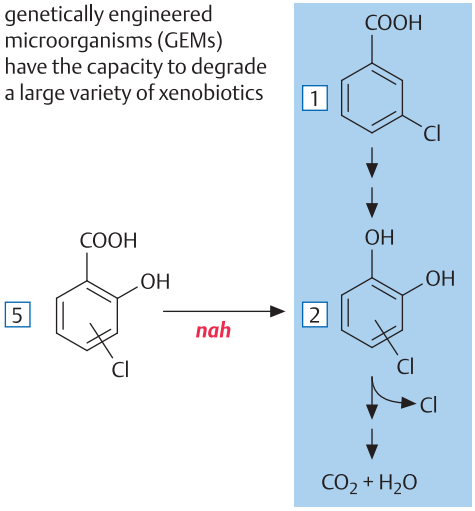




analysis → use/deposition of purified material ← analysis

Extension of degradation pathways

genetically engineered microorganisms (GEMs) have the capacity to degrade a large variety of xenobiotics



added genes:

- ***bph*** degradation of 3 to 4
- ***xyl*** degradation of 4 to 2
- ***nah*** degradation of 5 to 2

1 3-chlorobenzoate
2 chlorocatechol
3 chlorobiphenyl
4 4-chlorobenzoate
5 chlorosalicylate

wild-type *Pseudomonas* sp. 13: no degradation of 3, 4, 5

Microbial leaching, biofilms, and biocorrosion

General. Leaching of metals from low-grade ores through inoculation with Thiobacilli is carried out mainly in the USA, Mexico, and Australia: ca. 25 % of the global production of copper, ca. 10 % of uranium, and ca. 3 % of cobalt and nickel production are carried out by bioleaching.

Microbiology and physiology. Metal leaching is done with Thiobacilli, Gram-negative rods that are obligate chemolithotrophic organisms (→12) and that assimilate CO_2. They generate energy by oxidizing reduced sulfur compounds, such as sulfides, to sulfuric acid. Besides *T. thiooxidans*, *T. ferrooxidans* can oxidize not only reduced sulfur compounds, but also soluble Fe^{2+} salts. To synthesize 1 g of cell dry weight, *T. ferrooxidans* oxidizes 156 g-equivalents of Fe^{2+}. Both bacteria are well adapted to growth under acidic conditions and tolerate pH values down to 2.0. Sulfidic and oxidic ores such as pyrite (FeS_2), calchocite (CuS_2), covellite (CuS), sphalerite (ZnS), lead-, molybdenum-, and antimony sulfides (PbS, MoS_2, Sb_2S_3), cobalt and nickel sulfides (CoS, NiS) and also oxides such as pitchblende (UO_2) can be solublized by Thiobacilli. During the *direct bacterial leaching process*, Thiobacilli oxidize sulfide minerals directly according to the equation:

$$\text{metal sulfide} + 2\,O_2 \rightarrow \text{metal sulfate}$$

via several intermediary steps. In contrast, the action of Thiobacilli in *indirect bacterial leaching* is catalytic, in that it assists the geochemical oxidation of sulfide minerals to Me^{2+} and sulfide oxidation according to the equation:

$$\text{metal sulfide} + Fe_2(SO_4)_3 \rightarrow$$
$$\text{metal sulfate} + 2\,FeSO_4 + S^\circ$$

thus promoting oxidation at highly acidic pH. At pH 2–3, bacterial oxidation of Fe^{2+} is about 10^5–10^6-fold higher than chemical Fe^{2+} oxidation. Due to the complex compositions of ores, both types of leaching in field applications may overlap. Heterotrophic microorganisms are also considered for bioleaching. Thus, *Aspergillus* or *Penicillium* strains which form strong chelating agents such as citric or gluconic acid bind metal ions as water-soluble complexes. However, organic C-sources are required for growth of these organisms.

Technology. For high efficiency, the following parameters must be met or optimized: the chemical composition and mesh size of the mineral, the mineral nutrient, very low pH, positive redox potential, temperatures around 30 °C, and a good supply of O_2. The technical process may be carried out *in situ* at mines, in ore piles, or in tanks. Alternatively, abandoned mine tunnels may be flooded for *in situ* leaching. The technology is most advanced for tank leaching, which competes best with pyrometallurgic processes if highly disperse concentrates of the metal predominate in the mineral and if environmental aspects come into play.

Biofilms and biocorrosion. Biofilms form when bacteria adhere to surfaces in aqueous environments and begin to secrete biopolymers that can anchor them to materials such as metals or tissues. Usually, a biofilm consists of many species of bacteria, as well as fungi, algae, protozoa, debris, and corrosion products. In the microbial corrosion of metallic iron, (Fe^0) is subject to "anaerobic oxidation" to FeS, catalyzed by anaerobic sulfate reducers such as *Desulfovibrio vulgaris* according to the equation:

$$4\,Fe + SO_4^{2-} + 2\,H_2O \rightarrow FeS + 3\,Fe(OH)_2$$

Iron is oxidized under anaerobic conditions according to the equation:

$$4\,Fe + 8\,H^+ \rightarrow 4\,Fe^{2+} + 4\,H_2$$

and the H_2 layer produced during this process protects the metal from further oxidation. In the presence of sulfate, however, *Desulfovibrio* reduces sulfate according to the equation:

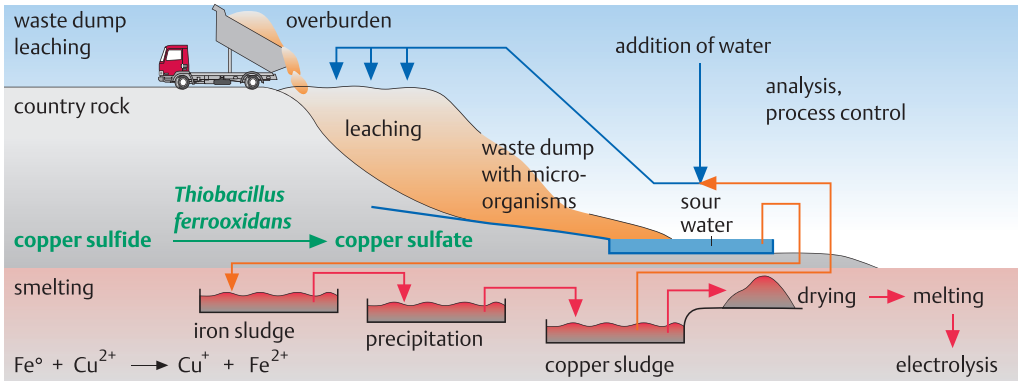
$$4\,H_2 + SO_4^{2-} \rightarrow H_2S + 2\,H_2O + 2\,OH^-$$

leading to further corrosion of Fe° by the precipitation of iron sulfate and iron hydroxyde:

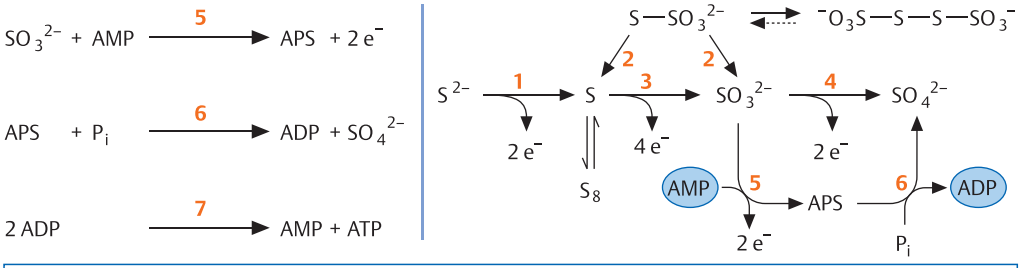
$$4\,Fe^{2+} + H_2S + 2\,OH^- + 4\,H_2O \rightarrow$$
$$FeS + 3\,Fe(OH)_2 + 6\,H^+$$

Other anaerobic microorganisms such as *Acidithiobacillus thiooxidans*, *Thiobacillus thioparus* or *Thiobacillus concretivorus* have also been found to participate in corrosion processes, because they can oxidize hydrogen sulfide – which is formed in anaerobic metabolism – to sulfuric acid. They also form biofilms. The damage to iron pipes due to biocorrosion is in the order of billions of €.

Microbial leaching



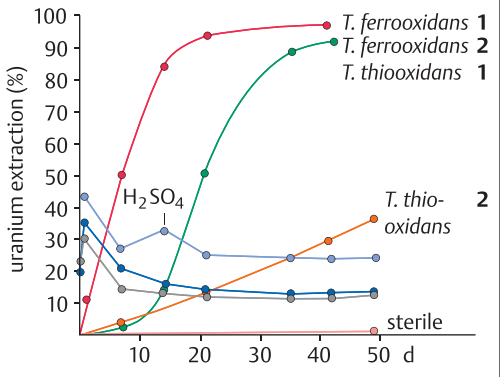
Biochemical reaction steps



AMP adenosine monophosphate
ADP adenosine diphosphate
APS adenosine phosphosulfate

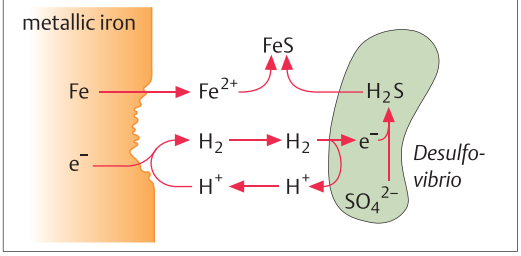
1 sulfide oxidase
2 rhodanese (thiosulfate hydrolase)
3 sulfur oxidase
4 sulfite oxidase
5 APS reductase
6 sulfate adenylyl transferase (ADP sulfurylase)
7 adenylate kinase

Kinetics of uranium leaching

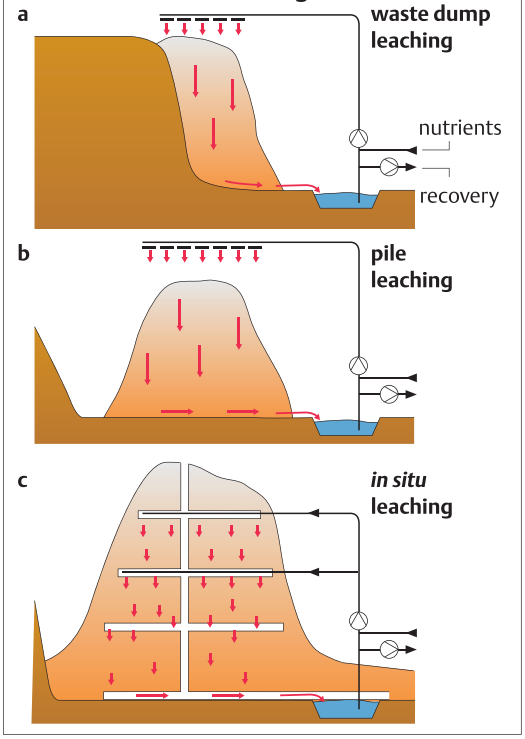


shake flash cultures, 5 % uranium ore, **1**, **2**: experiments with different particle size

Biocorrosion



Variations of ore leaching



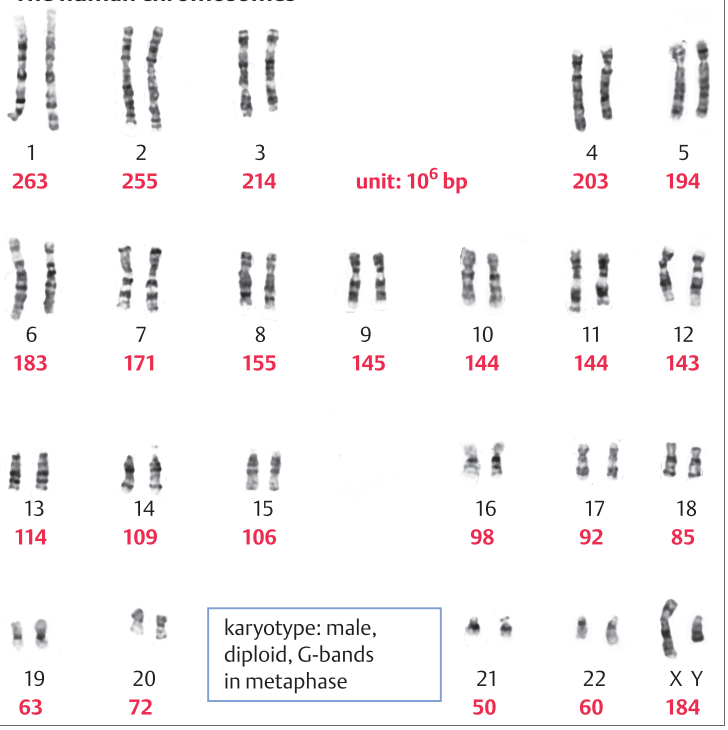
The human genome

General. Except for genome sequencing or gene therapy, genetic experiments with humans are neither legal nor desirable. However, because the pattern of population growth means that 6% of all people who ever lived on earth are living now, an extensive gene pool of humans is available. Genetic analyses of inheritance within families have led, over many decades of research, to a chromosomal map in which, even before the advent of human genomics, many hundreds of genetic diseases were roughly localized. Unambiguous localization of a disease in terms of the DNA sequence is, however, still rare. Within the human genome, which is ca. 3 billion (10^9) bp in size and distributed over 23 chromosomes (haploid set), only a few percent of the entire amount of DNA is sequence information coding for proteins. Most of the genome sequence consists of repetitive elements whose function is still unknown. Single nucleotide polymorphisms (SNPs) in genomic and mitochondrial DNA of contemporary humans allowed for an analysis of human migration over the last 50,000 years ("Genographic Project") (→298).

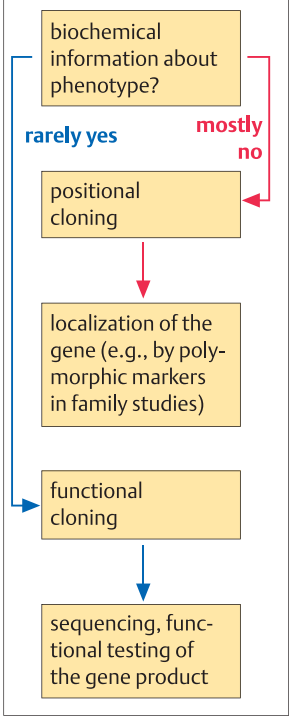
Genetic mapping. Key phenotypic markers in humans are inherited diseases, which in a few cases have been mapped by methods of human genetics (family analyses, chromosome analyses, functional and positional cloning of single genes). The Centre d'Etude du Polymorphisme Humain (CEPH) (→298), founded in 1984 in Paris, maintains tissue cultures of ca. 100 families from 3 generations, which on average comprise 4 grandparents, 2 parents, and an average of 8 children. In addition, populations with little mobility (e. g., Iceland, Tasmania) have been investigated in the hope of correlating genetic polymorphisms with the phenotypes of inherited diseases, leading to a functional gene map. Useful markers for such endeavors are microsatellites: they occur in high frequency and at the same time are highly variable, resulting in a probability of 70% that each individual is heterozygous for any given microsatellite. As a result, gene loci coupled to a microsatellite can be individually traced. In 1994, it became possible for the first time to generate a human genetic map with one marker per ca. 600 000 bp, combining the analysis of 291 meioses (304 individuals from 20 families of the CEPH collection) with the positions of 2335 microsatellites. In the meantime this map has been further refined.

Human genome sequencing. This large-scale project was begun in 1990 by an international, publicly funded consortium. Its strategy was based on contig sequencing of overlapping clones. Individual chromosomes were tagged with a fluorescent marker and separated by fluorescence-activated cell sorting (FACS) (→84). Their DNA was cut and transferred into ca. 300,000 BAC clones (→68); clones with overlapping DNA sequences were identified by restriction mapping, chromosome walking, and sequence-tagged sites (STS); their DNA was end-sequenced, after subcloning, in both directions (to eliminate sequencing errors); and the overall sequence was computed using genetic maps for validation (→72). Since 1996, this procedure was complemented by the sequencing of ca. 50,000 non-redundant expressed sequence tags (EST). Beginning in 1998, a private company, Celera, started to compete with the public project, using a shotgun approach. To this end, human DNA was cut into 60 million fragments of ca. 2000 bp and 10 million sequences of ca. 10,000 bp, which were sequenced from both ends in lengths of 500 bp at a time (the 10,000-bp fragments were sequenced to ensure that repetitive sequences up to 5 kbp long were not wrongly assigned). The total length of DNA sequenced by this approach was ca. 35 billion bp, corresponding to a roughly 12-fold redundancy. Both projects have been completed. A rough draft of the human genome sequence was published by mid-2000, and a final sequence in 2004. Surprisingly, only 20,500 genes were found, the same range as in mice. However, the human genome shows more segmental sequence duplications than other mammals. In the past 10 years, high-throughput sequencing (HTS) has facilitated the sequencing of thousands of human genomes. A "human reference genome" was built based on 13 volunteers, and the focus has shifted to understanding individual and regional variations in genome sequence (→298). Their phenotypic significance in health and disease have now become a major area of research. Epigenetic patterns (→66) in human genomes are being investigated, and the diversity of primary gene products in different cell types has become a major objective of proteomics research (→314).

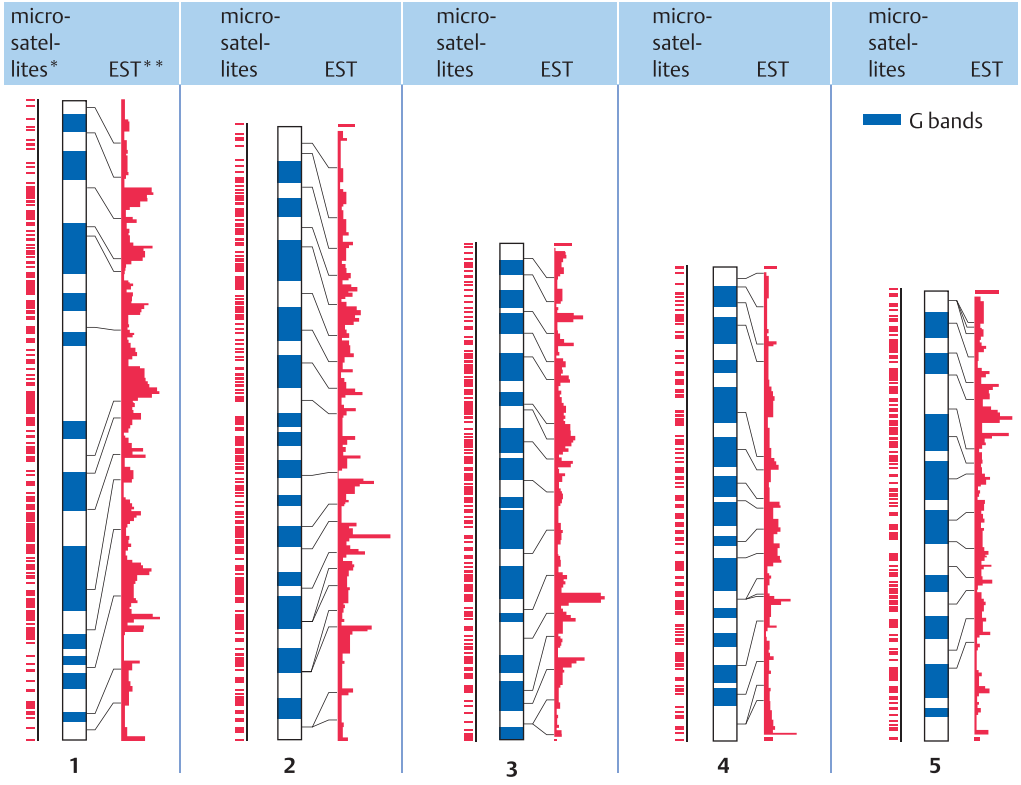
The human chromosomes



Mapping



Physical gene map with markers (chromosomes 1 – 5)



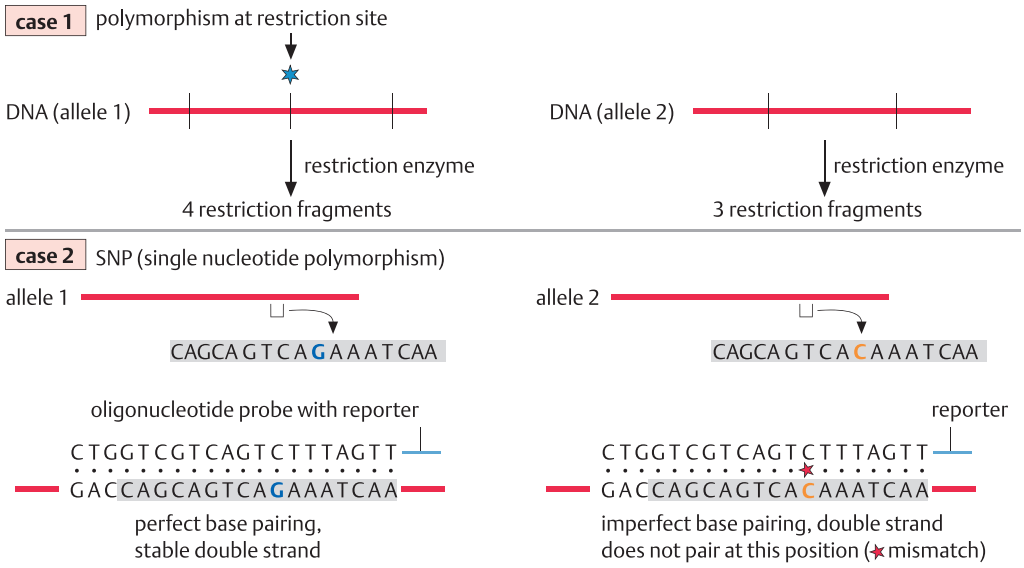
* 2 335 microsatellites from CEPH families (CEPH = Centre d'Etude du Polymorphisme Humain), cell bank from individuals of one family over 3 generations, **20 104 expressed sequence tags (EST)

Functional analysis of the human genome

General. After the first sequence of a human genome was published in 2004, further programs focused on human genomic diversity and its evolution. Thus, the International HapMap project (2005–2010) explored genomic variance in 269 individuals chosen from the USA, Japan, China and Nigeria, and the "1000 genomes project" (2007–2012), which was focused on frequent genetic variations in humans. Among many other related projects are the cancer genome project (since 2007), the human microbiome project (mainly on human intestinal microorganisms, since 2008) (→118), the human connectome project (molecular genetics of human neuronal diseases, since 2009), the Neanderthal genome project (genome draft 2010), or the chimpanzee and bonobo genome project (genome drafts 2005 and 2012). Their target is and was to understand the molecular physiology and pathophysiology of man and his immediate relatives on a genetic basis, and to use this knowledge for an improved therapy of monogenic and polygenic diseases. In a few cases, this was already successful, e. g., cystic fibrosis (→232), which can be diagnosed as a single nucleotide polymorphism (SNP), or in Leber's congenital amaurosis (LCA) where gene therapy was successful in several cases. In most cases, however, genetic diseases have polygenic origins, and they may evolve in a lifespan by a combination of inherited alleles, aging processes, nutritional deficiencies or epigenetic events (→66). Other procedures, which help to analyze the molecular biology of human diseases are 1) transcriptomics, by DNA arrays ("which genes are being transcribed in health and disease?") (→316), 2) proteomics (→314) ("which proteins are formed by the transcriptome, and how are they being processed?"), 3) the analysis of metabolism (→318), and 4) further procedures.

Comparative genetic diagnostics. Each human individual possesses many SNPs, distributed at a ratio of ca. 1 per 500 bases or a total of ca. 6 million over the whole genome. In the framework of the HapMap project, over 1 million SNPs each from European, Chinese, Japanese and Nigerian individuals were deposited in a database. From the "1000 Genomes" project, a genome map was generated in 2012 that locates genetic variations of 1092 individuals that are located in genomic regions believed to bear relevance for a number of diseases. Samples were obtained from Europeans, Africans, North and Latin Americans and North and Southern Asians. If a SNP is located within the recognition sequence of a restriction enzyme, it can be analyzed be RFLP (restriction fragment length polymorphism) as cutting of this DNA section by restriction enzymes (restriction analysis) leads to an altered fragment pattern. Most SNPs and many RFLPs are „silent" mutations, their phenotypic consequences, if there are any, are unknown. Those located in microsatellite regions have proven very useful for the genotyping of individuals, e. g. for parental testing and in criminal investigations. The correlation of functional gene sequences with monogenic diseases and variations such as hemophilia or brachydactyly (shortness of fingers and toes) is still very difficult, however, and impossible for polygenic phenotypes or diseases such as, e. g., the color of hair, the risk for cancer, or human intelligence. Analysis of human inheritance is limited to observation, but many useful discoveries have been made by comparison with animal genomes, which are amenable to experimentation. Thus, even phylogenetically remote living beings such as a fruitfly, a worm, a fish, and yeast (Drosophila, Caenorhabditis, zebra fish, Saccharomyces) exhibit gene functions that are related to those in humans and, unlike humans, can be subjected to genetic experiments. In this context, the mouse genome, which has also been sequenced (ca. 3.3 billion bp) is of particular value, because the genetic and physiological relationship between mice and men is high. Knockout mice (→270), where a specific gene or regulatory element has been impaired ("Alzheimer's" mouse, SCID-mouse), thus play a key role in fundamental research and in developing treatments. The results of such studies increase the chance of predicting, in the future, from a few individual gene sequences, the risk that an individual or his/her offspring will develop a certain disease. Of particular value to this end are transcription analysis of large gene clusters by DNA arrays, gradually displaced by high-throughput genome sequencing, and investigations of the proteome. No doubt these rapid developments will raise many new ethical questions (→336).

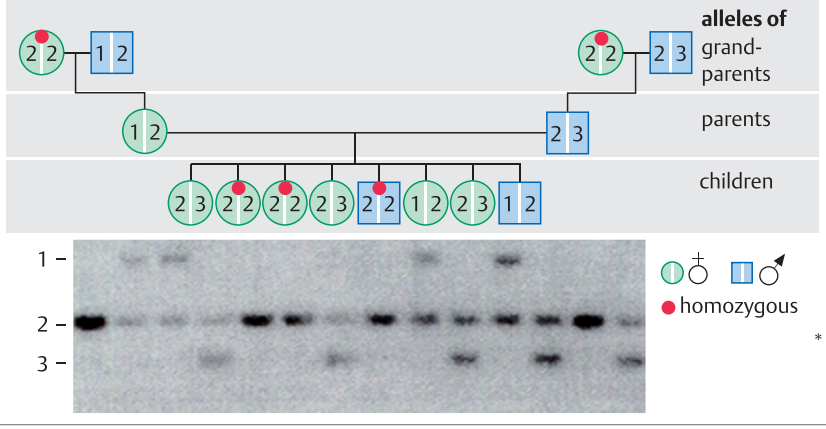
Polymorphisms and SNP analysis



Programs (selection)

program	Topic	Remarks	Duration	URL
CEPH Fondation Jean Dausset	CEPH families reference panel, YAC libraries of the human genome	genes associated with longevity, Crohn's and Alzheimer disease,asthma, cancer	1984–	www.cephb.fr/en/
international HapMap project	genetic variance in different human populations	based on 270 individuals from Nigeria, Japan, China and the USA	2005–2010	hapmap.ncbi.nlm.nih.gov
1000 Genomes	to find most genetic variants that have frequencies of at least 1% in the populations studied	genome sequences of about 2,500 samples from various populations at 4X coverage	2010–	www.1000genomes.org
human microbiome projects	genomes of all microorganisms living in association with the human body	reference set of 3000 microbial genomes established, national projects	2005–	hmpdacc.org/overview/about.php www.human-microbiome.org
cancer genome project	prepare a cancer genome atlas	includes specific cancer types	2006–	cancergenome.nih.gov

Inheritance of an RFLP pattern



lymphoblastoid cell cultures of a CEPH* family, from which an RFLP probe for a gene locus on chromosome 5 was constructed and used to show its inheritance pattern.

* CEPH = Centre d'Etude du Polymorphisme Humain

Pharmacogenomics, nutrigenomics

General. Each patient shows a different reaction towards medications – a reason to carry out comprehensive clinical studies on risks and side-effects of drugs towards an individual patient or specific patient groups. As genomic variations among individuals contribute greatly to such effects, the progress of genetic analysis (→302) and high-throughput sequencing (HTS) (→312) of human genomes has raised expectations that the reaction of a patient towards a drug can predicted: if all steps in the metabolism of a drug and its targets are known, individual differences which are genetically encoded in these targets might allow an estimate of drug efficacy (pharmacogenomics) and compatibility for an individual patient ("personalized medicine"). Moreover, the choice, the structure or formulation of a drug might be tailored to an individual patient leading to a patient-specific therapy. Similar views hold for nutrition, though the situation here is much more complex: not only is food a complex mixture of chemical substances in a complex matrix, but microorganisms in the digestive tract (the human "microbiome") (→118) interact in food digestion and metabolism.

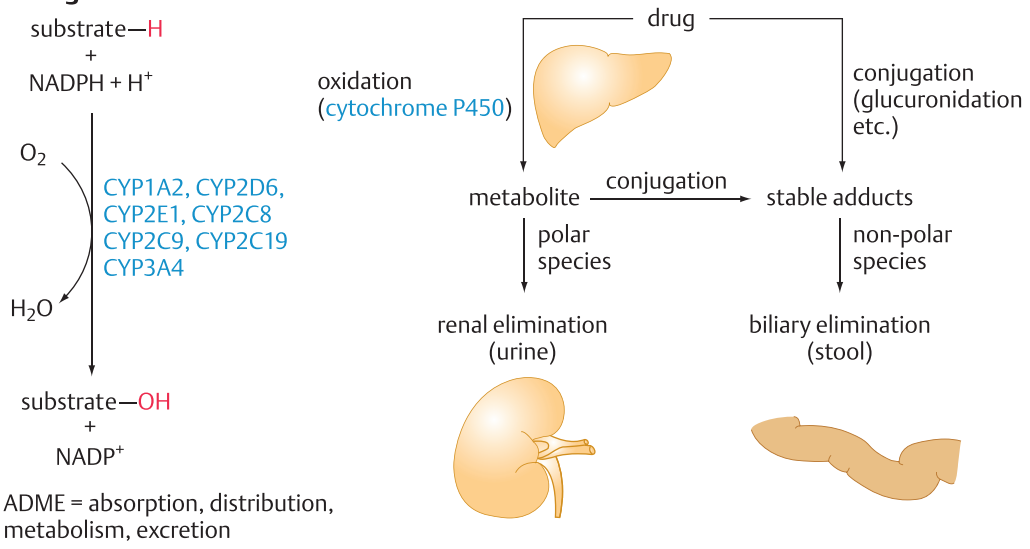
Drug metabolism (Pharmacokinetics). It has been long known that individual patients react towards the same drug in different ways. ~80 % of all drugs are "activated" (modified) via the cytochrome system of the liver, and important individual differences may already occur at this step. As an example, the monooxygenase cytochrome P4502D6 is one of the most important human enzymes for the activation and detoxification of xenobiotics (they are first hydroxylated and then coupled to hydrophilic metabolites such as glycine or D-glucuronic acid, rendering them water-soluble for excretion). If a patient is homozygous for this enzyme, encoded in chromosome 22 (homozygous: a functional allele from each parent), he or she is a "fast metabolizer." This is the normal case. Patients with duplications of this gene are extremely fast metabolizers (less of the drug reaches the target), heterozygous patients with only one parental allele are slow metabolizers (more of the drug reaches the target). Mutations of the enzyme may also lead to changes in activity. Individual differences in genetics thus effect pharmacokinetics (ADME (→310) = absorption, distribution, metabolism, excretion) and can lead to an overdosage or underdosage of drugs in their active form at their target; worst cases would be that the drug is either completely metabolized before it reaches its target, or not metabolized at all, resulting in excessive concentrations on target and toxic effects.

Drug action on target. Similar considerations apply to the reaction of a drug with its target, e. g., a receptor. Thus, Enalapril, a drug to treat hypertension and an inhibitor of the angiotensin-converting enzyme (ACE), acts more strongly on some mutants of this enzyme. Trastuzumab, a therapeutic antibody for the treatment of breast cancer, is active only in those 1/3 of all patients who overexpress in their tumor the Her2 receptor, the target protein. If nucleoside analogs such as 5-fluorouracil, 6-mercaptopurin and others are used for the chemotherapy of tumors, the patient's polymorphisms for dihydropyrimidine-dehydrogenase (DPD) and for variants of UDP-glucuronosyltransferase (UGT1A1) need to be checked in advance, since a deficiency in these enzymes leads to slow detoxification of these toxic drugs and severe side effects.

Companion diagnostics. Since knowledge about the impact of polymorphisms on the doctor's choice for treatment of a disease is rapidly increasing, genetic diagnosis of such polymorphisms is gaining momentum and has lead to the concept of "companion diagnostics" which help to analyze the specific genotype of a patient before prescribing a drug. Thus, a test for the presence of mutation V600 of the proto-oncogene B-Raf in patients with metastasizing melanoma (50 % of all cases) determines if an oral therapy with vemurafemib, a protein kinase inhibitor, has good prospects.

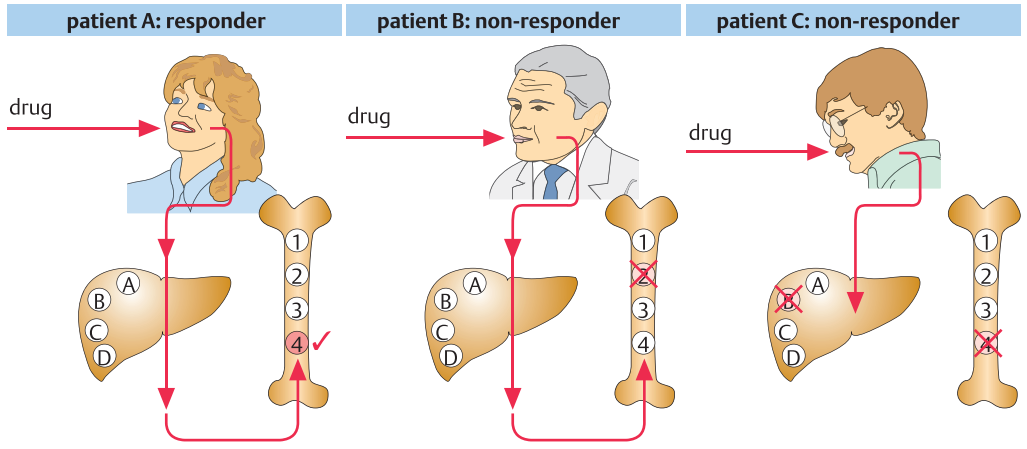
Nutrigenomics. This new field attempts to relate individual human polymorphisms with the activity of bioactive food components. This might eventually lead to optimized individual diets based on genetic information. Examples are the identification of polymorphisms in the leptin gene, which have been related to adiposis, and in the methylentetrahydrofolate reductase (MTHFR) gene, a key enzyme in folic acid metabolism whose malfunction leads to severe health problems.

Drug metabolism



ADME = absorption, distribution, metabolism, excretion

Pharmacogenomics



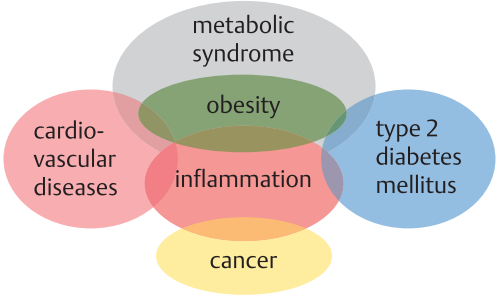
Ⓐ – Ⓓ metabolic steps for drug activation

① – ④ metabolic steps related to disease

Companion diagnostics

	companion diagnostic assay: does drug meet targets?	consequences for therapy
drug →	maybe	unclear outcome: weigh benefits and risks
drug →	yes	patient shows positive response to treatment
drug →	no	safety and efficacy questionable, stop treatment

Nutrigenomics



Interactions of

- genes involved in human metabolism
- bioactive food components
- genes in the human microbiome

DNA assays

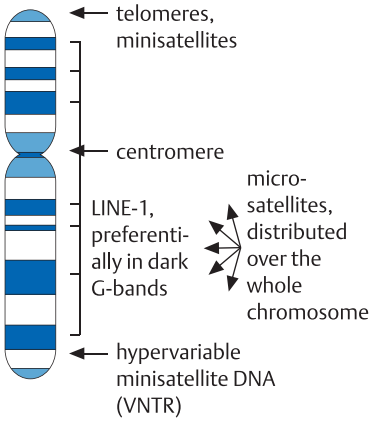
General. Hybridization events that occur between two polynucleotide strands (→38) (DNA-DNA, DNA-RNA, or RNA-RNA) through hydrogen bonding can be visualized by tagging one of the strands with a radioactive, fluorescent, or otherwise detectable marker (→84). Usually, the DNA or RNA to be analyzed is first amplified by PCR (→50). To minimize false-positive or false-negative hybridization events, a thorough protocol for sample preparation (samples may be as diverse as urine, blood, tissue, plant materials, or fossils), the choice of the single-stranded probe, and the hybridization conditions are of utmost importance. DNA assays are now used in many areas, e. g., for genotyping pathogenic microorganisms, for genetic analysis of diseases, for monitoring genetically modified plants or foods derived from them, for parental testing, and in criminal investigations. The market volume for DNA assays is growing rapidly and is already on the order of several billion US$.

Equipment. Standard procedures for DNA assays are 1) comparative gel electrophoresis (→54), 2) hybridization assays based on optical or electrochemical reporter groups (→84), and 3) DNA arrays (→316). In electrophoresis, the sample is analyzed in comparison to DNA of standard sizes. Denaturing gradient gel electrophoresis (DGGE) is a useful procedure for obtaining fingerprints of microbial communities, e. g., of the intestinal biota. For assays based on hybridization, reporter groups can be linked to a single-stranded polynucleotide by standard chemical procedures, e. g., by biotin-avidin/streptavidin linkers. Alternatively, as in the LightCycler™ assay (→50), an intercalating dye (SYBR Green™) is inserted into amplified DNA, allowing for real-time quantitative measurement of hybridization events. Another useful procedure is the TaqMan™ Assay, which is based on a "molecular beacon" (a fluorophore) attached to the end of the DNA probe. It interacts with a quencher at the other end of the probe. When a matching sequence is present in a sample that is subjected to PCR, the probe is amplified, the fluorophore is liberated, and the sample fluoresces. Due to the rapid technical progress, massive parallel sequencing (→312) is already competing with hybridization, such as in "liquid biopsy" of tumor DNA in blood indicating metastatic proliferation.

Genotyping is widely used in forensic and paternal tests (genetic fingerprinting) and is emerging as a powerful method for the analysis of human single-nucleotide polymorphisms (SNPs). In paternal testing, a small number of microsatellite sequences of a male are compared with the microsatellite sequences of his putative offspring or, in criminal investigations, a suspect's sequences are compared with a sample containing DNA traces left at the crime scene. The probability that two individuals share the same microsatellite pattern is extremely low (except for identical twins). SNP analysis of a patient is an emerging technique to determine, e. g., the individual safety and efficacy of drugs (pharmacogenomics) (→300). Nearly 113 million human SNPs had been entered into the SNP database of the US-based National Center for Biotechnology Information (NCBI) as of the end of 2014. SNP analysis is also used in plant and animal breeding to secure pedigrees (→264, 274). In medical diagnosis, DNA assays are used to identify gene sequences and their aberrations and also the presence of infectious organisms in a blood, liquor or urine sample. A 1μl blood sample is enough to demonstrate the presence of *Plasmodium falciparum*, the causative agent of malaria, in a patient. With DNA assays, food samples can be analyzed for the presence of extremely low levels of pathogenic microorganisms or genetically modified raw materials. Since genotyping assays generally start by amplifying the target DNA from the sample, the method implies that enough information is available about the sequence of the target DNA to prepare a synthetic probe that can hybridize with the target.

Prenatal screening is mostly based on ultrasound combined with biochemical markers. It is sufficiently precise to detect monogenic diseases in a human embryo in weeks 9–12 of pregnancy. Precision is raised by targeted sequencing of DNA, either isolated from an embryonic cell or from fetal DNA in maternal blood. The extension of this method to risk analysis for mono- and polygenic diseases in the preimplantation stage (preimplantation diagnostics, PID) has already been established in principle, but is hotly debated with regard to ethics (→336).

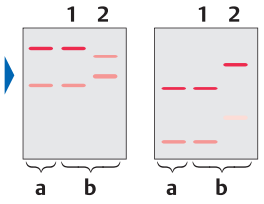
Forensic DNA analysis



genotyping of several gene loci

DNA fingerprinting in forensic medicine: DNA is isolated from a secured sample, several microsatellite-containing regions are amplified by PCR, and Southern blot analysis is carried out using a range of selective probes

2 probes each hybridizing with one gene locus



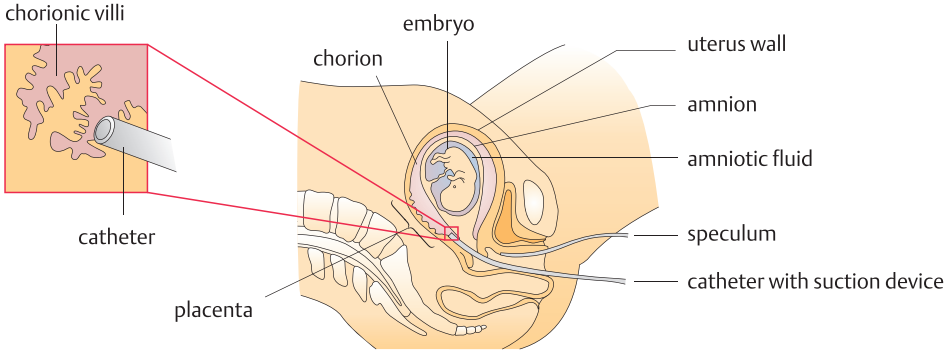
a DNA sample from place of crime
b DNA samples from two suspects 1 and 2; suspect 1 is indicted

microsatellites	< 1 kbp	repetitive $(AC)_n$ sequences, n = 10 – 50
minisatellites	1 – 30 kbp	many repeats in sequence, variable number of tandem repeats (VNTRs), subtelomeric location
macrosatellites	< 1000 kbp	usually in the centromeric or telomeric region, AT-rich
randomly inserted repetitive sequences	SINE: 100 – 500 bp LINE: 6 – 7 k bp	~ 20 % of the human genome; ~ Alu-family: 5 % ~ 10 % of the human genome

DNA analysis of infectious diseases (examples)

- human papilloma virus (HPV)
- herpes virus, e.g., cytomegalovirus (CMS)
- HIV virus
- *Mycobacterium tuberculosis*
- *Borrelia burgdorferi*
- *Helicobacter pylori*
- *Neisseria meningitidis*

Genetic screening



analysis of chorionic villi (weeks 9–12)
methods of analysis
chromosomal anomalies
biochemical tests
FISH-based methods
DNA tests

DNA tests for some monogenic diseases

disease (frequency)	**gene (position)**
sickle cell anemia (1:500; blacks)	(11 short)
cystic fibrosis (1:2 500)	CFTR (7q31)
hemophilia A and B (1:25 000)	(Xq27)
Huntingdon's disease (1:20 000)	IT15 (4p 16.3)
phenylketonuria (1:10 000)	PAH (12 long)
inherited form of breast cancer (1:200, females)	BRCA1 (17q21)

Gene therapy

General. Among the ca. 15,000 human diseases that have been reported, over 90 % cannot be cured by rational therapy. Most of them originate in inherited or acquired genetic defects. Gene therapy is the attempt to replace genes having impaired function by healthy genes. As of the end of 2013, over 1700 gene therapeutic protocols have been used for the treatment of thousands of patients, >60 % in the USA. Gene therapy directed towards somatic human cells is generally accepted, but gene transfer into human sperm or egg cells (germ line), which would lead to inheritable new properties of the recipient's offspring, is not and is subject to a moratorium (→336). It is practical to distinguish gene therapy *ex vivo*, in which human cells are multiplied (expanded) and transformed outside the human body before they are re-infused, and *in vivo* gene therapy, which is direct therapy of patients with genetic material.

General concepts. For most diseases, the genetic basis is still largely unknown. But even gene therapy of monogenic diseases having a known genetic cause meets with great difficulty: the immunological barriers of the body and the cellular control mechanisms directed towards foreign nucleic acids must be overcome. Presently, experimentation is focused on 1) recombination of defective gene sequences with added cDNA of correct sequence, 2) silencing of genes via antisense or interfering RNA, and 3) repair of defective DNA sequences with RNA-DNA chimeras. Preferred vectors include retroviruses (ca. 20 % of all protocols), adenoviruses (ca. 23 %), vaccinia viruses (ca. 8 %) (→6), or cationic liposomes for lipofection (ca. 12 %) (→58). Liposomes permit larger cDNA fragments to be transferred; however, the space available for foreign DNA inserts in a viral capsid is small – between 4 kbp (adenovirus) and 30 kbp (Herpes virus). Liposomes (→34) can be administered as an aerosol through the respiratory tract, reaching cells via endocytosis. Viral vectors are usually applied subcutaneously, into the muscle, or directly into a tumor. The transfer of the patient's own bone marrow after modification with DNA (autologous cell therapy) and the direct application of DNA have also been described.

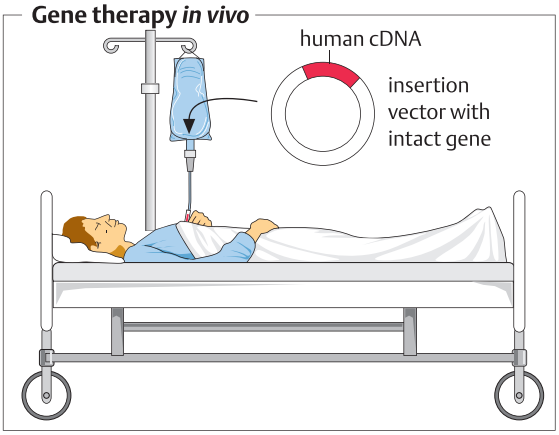
Individual protocols. About ⅔ of all gene therapeutic protocols are focused on tumor treatments. To this end, a tumor-suppressor gene such as BRCA1 or p53, a cytokine gene such as IL-2, a histocompatibility antigen such as HLA-B7, or a so-called suicide gene is administered. Protocols for the treatment of monogenic diseases often concern human severe combined immunodeficiency (SCID), which is caused by defects in the gene coding for adenosine deaminase (ADA). Gene therapy of infectious diseases (vaccination with naked DNA) is another major field of research.

Gene transfer ex vivo uses protocols successfully established for bone marrow transplantation. A preferred cell type is the hematopoietic stem cell, the precursor cell for all cells of the immune and blood systems. If it should become possible to replace a genetic deficiency in these yet undifferentiated cells outside the body by gene transfer, these transgenic cells should, after transfusion into the donor, differentiate into "healthy" immune and blood cells. The progress made with induced pluripotent stem cells (iPS cells) derived from the patient is strongly stimulating this area of research. The corrective gene can be introduced into the iPS cells using, for instance, a human artificial chromosome. Successful model experiments on iPS-induced cell (→306) cultures or on mice have already been reported for, e. g., the correction of Duchenne muscular dystrophy, spinal muscular atrophy or various immune deficiencies.

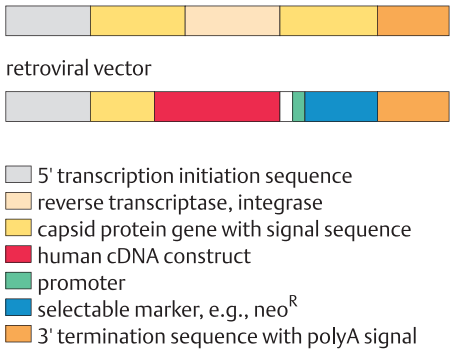
Gene transfer in vivo. In some of the many attempts to replace defective genes by transfecting a patient with vectors, liposomes, or DNA, at least part of the target cells were transformed. Sometimes, the clinical condition of patients suffering from, e. g. Parkinsons's disease, hemophilia, acute or chronic lymphocyte leukemia, etc. improved significantly.

Conclusion. Although gene therapy of monocausal diseases in humans is possible in principle, many unsolved problems remain. By 2013, a total of 3 casualties occurred, one of which was due to the adenovirus vector employed. However, a first drug for gene therapy has also been registered both in the USA and Europe: Glybera™, a gene coding for human lipoprotein lipase (LPL) and linked to an adeno-associated virus, can be applied to treat lipoprotein-lipase deficiencies (LPLD).

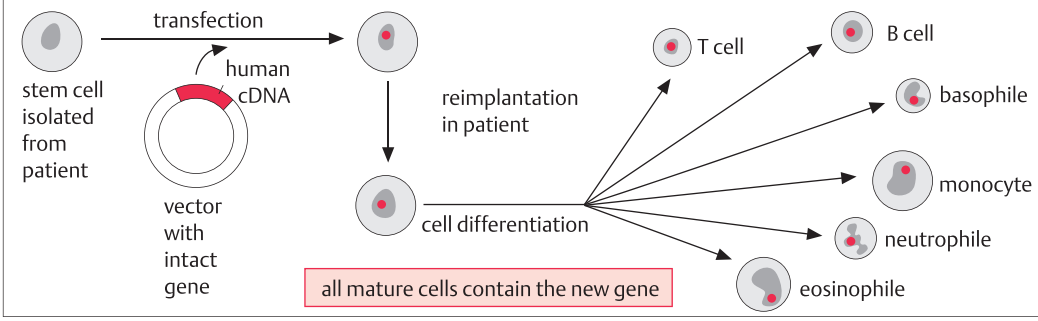
Gene therapy *in vivo*



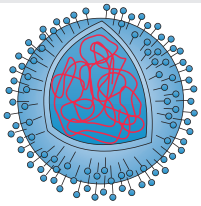
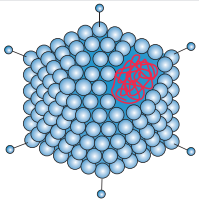
Retroviral vector



Gene therapy *ex vivo*, e.g., in stem cells



Vectors for gene therapy

retroviruses	adenoviruses	adeno-associated viruses	liposomes	naked DNA
				
advantage stable insertion into genome **disadvantage** statistical insertion, only dividing cells are infected	**advantage** incorporate large DNA segments **disadvantage** insertion in genome is unstable	**advantage** stable insertion into genome **disadvantage** low capacity for foreign DNA	**advantage** low infection risk **disadvantage** low efficiency	**advantage** low infection risk **disadvantage** low efficiency and stability

Experiments on gene therapy (2014)

disease	examples/transferred genes
cancer protocols: 1331 protocols	histocompatibility antigens, tumor-suppressor genes, suicide genes, IL-2, IL-7 and IL-12
monogenic diseases: 188 protocols	SCID ADA gene, cystic fibrosis, factor IX, chronic granulomatosis
infectious diseases, mostly AIDS: 170 protocols	transgenic T-lymphocytes, DNA vaccines
other diseases: 387 protocols	VEGF121 (atheriosclerosis), rheumatoid arthritis

Gene Therapy Clinical Trials Worldwide, http://www.abedia.com/wiley/index.html

Induced pluripotent stem cells (iPS)

General. In 2007, a Kyoto team led by Shin'ya Yamanaka successfully regenerated pluripotent stem cells (→78) from differentiated mouse cells by expressing a specific choice of genes. This result was later confirmed using human somatic cells. These experiments laid the foundation to generate pluripotent stem cells from the body's own somatic cells, and to apply them in a personalized cell therapy as well as for personalized drug testing. This procedure raises fewer medical and ethical questions than cell therapies based on harvesting embryonic stem cells (EMS) from umbilical cords, and thus has quickly been translated into clinical research on a global scale. At present (2014), several clinical tests for the therapy of coronary, eye, nerve and other diseases are already taking place. Yamanaka received the Nobel prize for medicine in 2012.

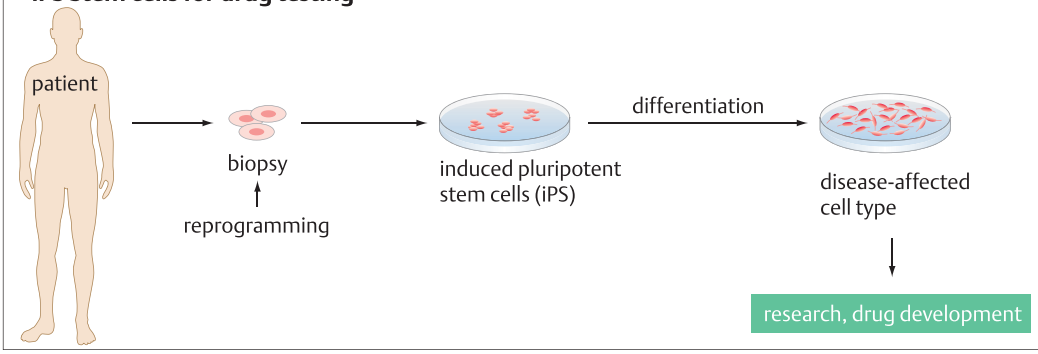
Methods. Yamanaka did his crucial experiments using mouse fibroblasts, which were transfected by a retrovirus carrying a choice of genes coding for transcription factors. He obtained pluripotent stem cells if the following four genes were introduced: 1) c-Myc, a transcription factor which bolsters up the expression of all active genes of a cell, 2) Klf-4, a transcription factor related to specific properties of stem cells, 3) Oct-4, a transcription factor involved in basic embryonic development, and 5) Sox-2, a transcription factor essential for the self-renewal of undifferentiated embryonic stem cells. A disadvantage of this first concept was the risk that tumor cells might develop, as c-Myc is a proto-oncogene, or that the retroviruses used for transfection might integrate into the genome of the mammalian receptor cells. The latter risk was minimized by using an adenovirus instead of a retrovirus for transfection. Later, it was found that both risks can be eliminated if receptor cells are directly transformed with the recombinant transcription factors (protein-induced pluripotent stem cells, PiPS). It also became possible to transform somatic human cells into pluripotent stem cells by synthetic modified RNA (RNA-induced pluripotent stem cells, RiPS). Today (2014), it is possible to generate iPS cells with almost no limit by an appropriate choice of somatic cells, transcription factors and transformation methods. Transformation yields, however, are still very low (< 1 %). Generation from human renal epithelial cells harvested from urine gave higher yields (up to 4 %) and good differentiation potential. Fat tissue is a good source for stem cells. Culturing in chemically defined media was enough to differentiate them into cells with the functional properties of hepatocytes.

Properties and risks. iPS cells and embryonal stem cells (ESC) (→78) have quite similar properties. Their behavior in differentiation, gene expression and methylation (epigenetics) (→66) is similar. Compared to ESC, iPS are artificially modified by transcription factors. As some of these factors are known to be involved in the formation of cancer cells, the safety of iPS-derived somatic cells is not yet resolved.

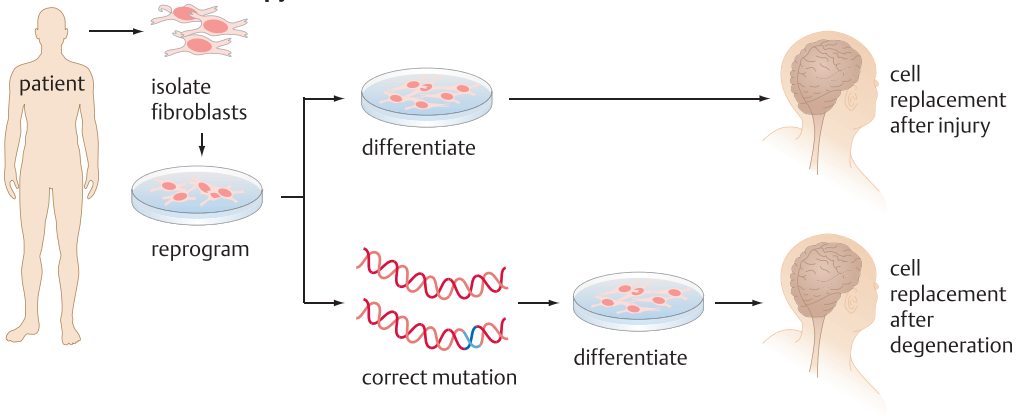
Expansion of iPS cells. iPS cells are adherent cells. They are usually grown on surfaces to which growth-promoting substances such as cadherin, an adhesion protein, are supplied. Complex but serum-free media are used for growth. They contain glucose, amino acids, γ-aminobutyric acid, inorganic salts, and recombinant proteins such as growth factors. After expansion, the cells are harvested from the surface by soft detachment, using, e. g., EDTA. The whole process can be carried out within robot-assisted work stations.

Applications. The two most important applications are in regenerative medicine (→308) and in the personalized testing of drugs. Both applications are expected to lead to innovative and large markets. In regenerative medicine, clinical studies are already being carried out with individualized iPS cells. In Japan, human patients with macular degeneration are treated with retinal pigment epithelium; patients suffering from Parkinson's disease are being treated with dopamine neuron progenitor cells; and patients with heart insufficiency are being treated with cardiomyocytes. To this end, the patient's cell were transformed into iPS cells, expanded in so-called cell sheets and implanted into the patient. Regeneration of blood cells such as T-cells involved in immune response are also being studied. For drug testing, several human cell types such as hepatocytes are already commercially available, and individualized formats are under development, a big step forward in personalized drug testing.

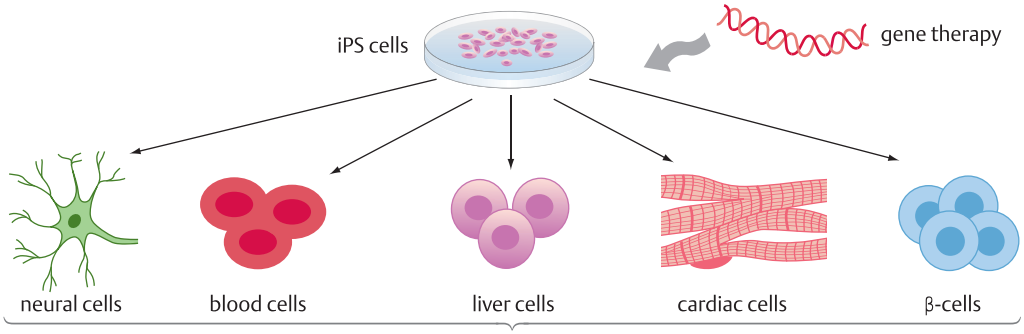
iPS Stem cells for drug testing



iPS Stem cells for therapy



Functional cardiomyocetes regenerated from iPS cells



Applications of iPS technology (selection)

target cells	starting cells	group, year
3D kidney structures (glomeruli, kidney tubuli)	various progenitor cells such as fibroblasts	Kumamoto University, 2013
human erythrocytes	various progenitor cells	CiRA, Kyoto University, 2013
human blood platlets	progenitor cells	Megakaryon Corp., Tokyo 2013
retinal pigment epithelial cells	progenitor cells	RIKEN, Kobe, 2014
human hepatocytes	progenitor cells	ReproCell Co., Tokyo, 2011
human cardiomyocytes	progenitor cells	Axiogenesis, Cologne 2012

Tissue Engineering

General. Tissue engineering aims at the functional regeneration of tissues through implantation of tissue cultured *in vitro*. Initial efforts on skin replacement for treating burns have been followed after several decades by tissue-engineered bone, blood vessels, liver, cartilage, ligaments, cornea, muscle, and nerve cells. Biocompatible matrices and omnipotent stem cells play an important role in this research.

Surgical approaches. In *autografting*, tissue is transplanted in the same patient from one location to another. For example, leg veins are grafted into coronary bypasses (ca. 400,000 in the US in 2010) or bone grafts from the hip are turned into a spine segment. Although this procedure does not pose immunological problems, it is costly and painful. Artificial materials and devices are widely used in, for example, artificial valves, prosthetic hips and knees, or breast implants.

Scaffold-guided tissue regeneration. Extracellular matrices (ECM) are the form- and shape-determining parts of the human body, for example, the composite fibrous network of collagen of bone. Artificial scaffolds such as ceramics, collagen tubes, or synthetic materials (films, membranes, sponges, beads) have been used for seeding donor cells. In the presence of appropriate cell growth factors (→238), the donor cells may grow on these scaffolds, yielding artificial tissue that can be implanted into a patient. Nourishing the interior cells that have grown into such scaffolds requires the formation of capillary vessels (angiogenesis) – a process that at present is not sufficiently mastered. CAD/CAM methods have been employed to fabricate arbitrarily complex scaffold shapes from CAD models (solid freeform fabrication, SFF). 3D Printing has opened an attractive new route to the production of body parts. Thus, a complete ear was generated from a polymer mold and tissues by three steps: 1) generating a 3D scan of the ear, 2) preparing a 3D hollow polymer form from these data using a 3D printer, and 3) filling this form with collagen and cartilage cells.

3D cell cultures. Various types of progenitor and primary cells can be co-cultured, forming a more complex tissue equivalent. For example, by growing (expanding) epithelial cells in the presence of keratinocytes, an artificial epidermis was produced. A 3D skin equivalent was obtained by co-culturing dermal fibroblasts, embedded in a collagen matrix, with keratinocytes, forming a cornified epidermis.

Stem cells. (→306) Pluripotent human embryonic stem cells have a major potential in tissue engineering, because they can develop into a wide range of different cell types, depending on the physical and chemical environment during growth and the cell growth factors present. Adult bone marrow is a traditional source of adult stem cells, with little ethical reservations. Stem cells can also be harvested from umbilical cord blood of newborn children and deep-frozen, to be used for the generation of tissues later in their lives. iPS technology has opened a broad avenue for the future generation of a patient's own tissues. Adipose tissue is rich in stem cells; adipocyte-derived stem cells have been shown to be turned into hepatocyte-resembling cells using defined chemical culture conditions.

Applications. Engineered tissues are widely used to 1) prepare test systems based on human cell culture, in the pharmaceutical and cosmetic industries; thus, for testing the irritancy potential of substances or for drug development, 3D skin and cornea, with cells of human origin, are used as a replacement for animal tests, providing more reliable results, and for 2) transplantation. Artificial human skin has been commercially available for some time. As an example, Dermagraft® is an engineered human dermal tissue composed of a bioabsorbable scaffold and human fibroblasts. It is used to treat burns and to support wound healing in different types of ulcers (diabetic foot ulcers, venous ulcers and pressure ulcers) by minimizing infections and retaining fluids until a sufficient amount of the patient's own skin is available for autografting. Mesenchymal stem cells residing in the adult bone marrow have been induced by cell growth factors to differentiate into chondrocytes, leading to cartilage tissue. Traumatic cartilage defects are treated by autologous cartilage transplantation (ACT), a method in which chondrocytes are injected into the injured area under a newly sutured periost flap. Replacement of bone by grafted bone tissue is widely used in regenerative medicine, and artificial nerve grafts or nerve guidance channels are being developed for nerve regeneration.

Targets in tissue engineering

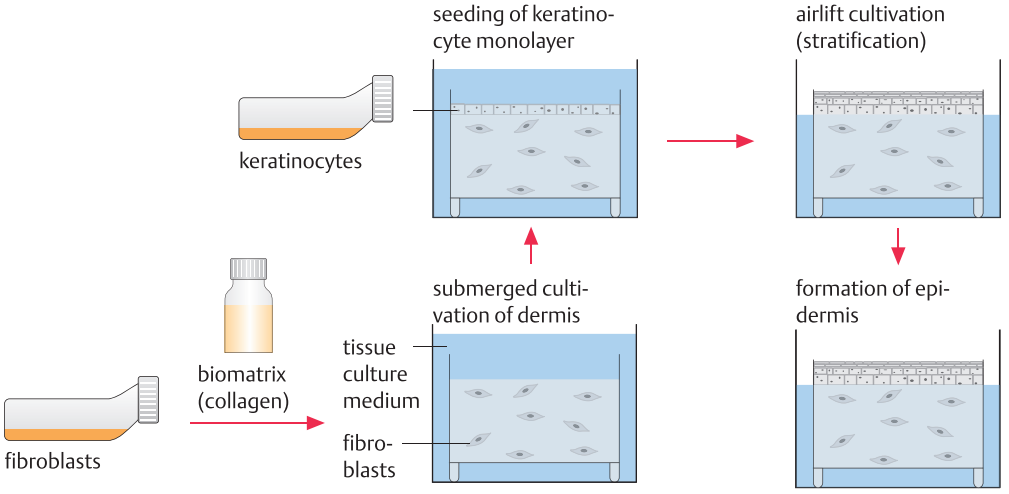
stem cell-based tissue engineering		non stem cell-based tissue engineering	
blood vessels	liver	bladder	meniscus
bone	pancreas	cartilage (ears, nose and joints)*	oral mucosa
cartilage	nervous tissue	heart valves	salivary gland
cornea	skeletal muscle	intestine	trachea
dentin	skin	kidney	ureter
heart muscle			urethra

*in clinical trials or clinical observational studies

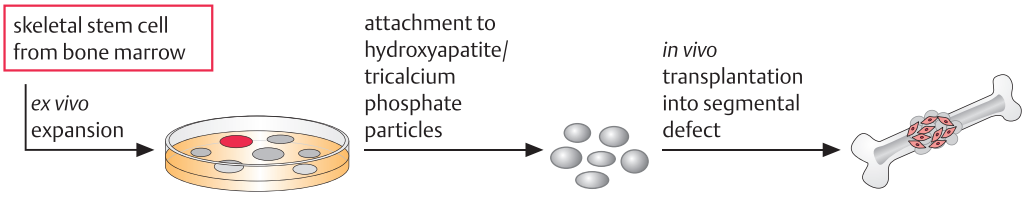
Some tissue engineering companies

USA	Advanced Tissue Sciences	human tissues and organs, e.g., skin, cartilage, bones, liver
USA	Curis	wound healing, regenerative medicine
USA	MatTek	skin test systems
Germany	BioTissue Technologies	skin, cartilage
Netherlands	IsoTis	skin, cartilage, bones
France	SkinEthic Laboratories	skin test systems

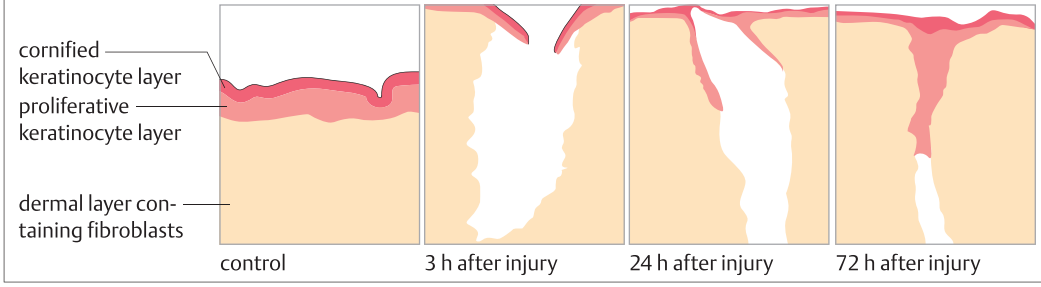
Generation of a 3-dimensional skin equivalent



Manufacture



"Wound healing" using artificial skin



Drug screening

General. New drugs and agrochemicals have traditionally been searched for by trial-and-error methods. Recent insights in the life sciences and the technical concepts of genetic engineering have replaced these methods by more rational approaches. They are based on the premise that drugs act on one or several targets within the human organism. Frequent targets are enzymes, receptors, or ion channels (→26, 34). The targets of agricultural agents may also be plant proteins that are involved in photosynthesis. In consequence of developments in genomics and proteomics, the functions of these targets are now much better understood. Targets can be prepared in sufficient quantities by recombinant technology, and their interaction with natural or synthetic compounds can be experimentally investigated. A main objective of this approach is to identify lead structures that interact with a given target. They are then modified by chemical synthesis, using the knowledge of medicinal chemistry on the effect of structural elements on ADME (→300) (absorption, distribution, metabolism, excretion), to synthesize novel, tailor-made, highly efficient drugs, or agrochemicals.

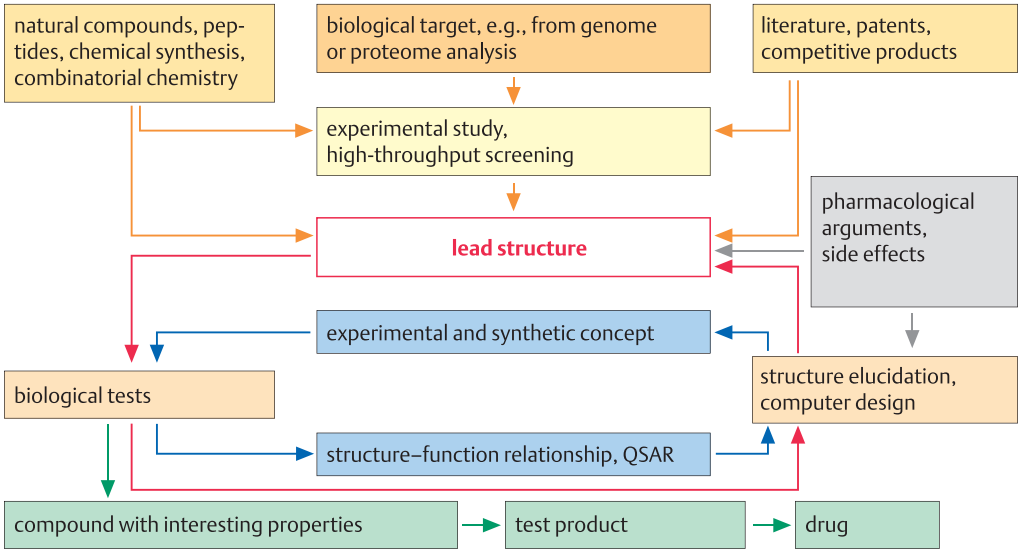
Identification and preparation of targets. The identification of targets is still difficult, especially for diseases of multigenic origin (the usual situation). By combining studies of inheritance patterns, for example (→296), in genetically isolated populations such as those in Iceland or Tasmania, with SNP, allele, or proteome analysis (→314) of patients suffering from a disease, valuable progress has been made recently. Once defects on an enzyme, an ion channel, a receptor – or a combination of them – have been identified as the putative cause of a disease, animal experiments are used to validate the hypothesis by knockout or RNA_i experiments (→64). If this approach leads to the validation of a target, it can be used in a high-throughput assay for screening new drugs. If, e. g., a G-coupled receptor has been identified as a target, it is expressed in the membranes of mouse fibroblast cells, while firefly luciferase is coexpressed as a reporter enzyme in the cells' cytoplasm. If a ligand binds to the receptor, signal transduction via cAMP-responsive elements raises the level of intracellular cAMP. Once luciferin, the luciferase substrate, is added to such cells, ligand binding can now be quantified by luciferase activity, which depends on cAMP. Binding events have also been detected *via* reporter genes, Ca^{2+} mobilization or cell impedance measurements (→84).

High-throughput screening. A widely distributed system in industry entails robot-assisted screening of large chemical libraries (100,000 or more substances) for interaction with a drug target. Combinatorial chemistry can increase the size of such libraries nearly without limit. A decisive factor for success is the availability of a rapid, valid assay for interaction between drug and target. Assays are usually carried out in 96-well, sometimes 384-well microtiterplates. Silicon-wafer technology in combination with confocal laser spectroscopy has further reduced the scale to nanoliter volumes, allowing the assay of >100,000 chemicals per day.

Rational drug design. Although the x-ray structures of most drug targets are still unknown, scientists have long attempted to postulate their drug-binding properties by systematic measurement of their binding of chemical analogues, leading to quantitative structure–activity relationships (QSAR). Based on a model of the binding-site structure derived from such analyses, they attempt to optimize the structure of a new drug by using molecular modeling and other computational approaches.

Future aspects. In the pharmaceutical industry, only one new drug reaches the market for each ca. 50 000 chemical substances investigated. Between 2008 and 2012, only ~20 new chemical entities (NCE) per year were introduced as new drugs into the market. In view of the many incurable diseases, and also for economic reasons, the global pharmaceutical industry attempts to increase its success rate by using target-based screening. Since a considerable number of putative drugs fail during clinical tests when human diversity comes into play, pharmacogenomics based on the analysis of individual disease-linked polymorphisms have begun to play an increasingly important role. If deviations in individual targets can be identified prior to therapy, e. g., through "companion diagnostics" (→300) in a DNA assay, side effects of drugs could be reduced or even be eliminated by fitting prescriptions to an individual patient ("personalized medicine") (→300).

Drug screening

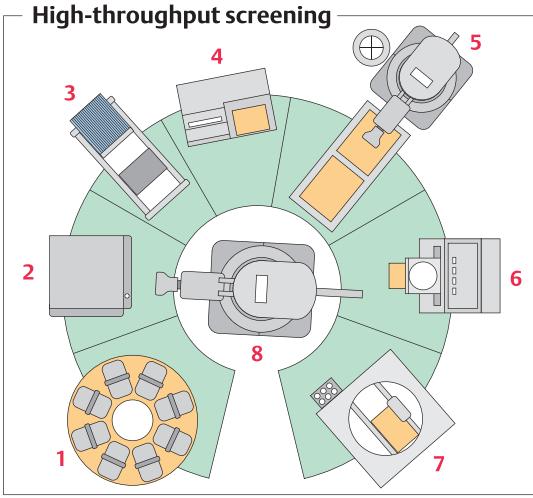


Human drug targets

	number*	example	application
enzymes	8 000	acetylcholinesterase	Alzheimer drug
receptors	15 000	serotonin receptor	schizophrenia
ion channels	3 000	Ca ion channel	old-age diseases

*man, estimates

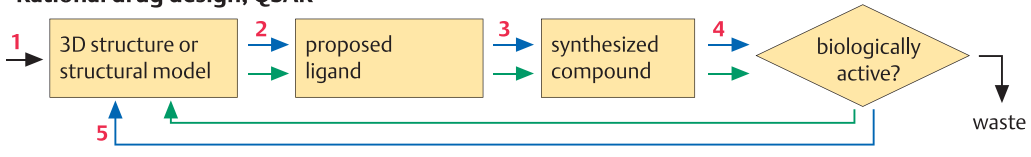
High-throughput screening



1 target cells in microtiter plates
2 CO_2 incubator*
3 addition of drug candidates and reagents
4 microtiter plate readers
5 parallel pipetting module
6 wash station for microtiter plates
7 centrifuge for microtiter plates
8 robotic arm that can approach stations 1 – 7 at random

* only required if human or animal cell cultures are used for expression of targets

Rational drug design, QSAR



1 structure elucidation or structural model of protein
2 design of ligands using structural model
3 chemical synthesis of ligands
4 biological test (assay)
5 structure or structural model of protein–ligand complex

→ cycle 1
→ cycle 2 and further cycles for optimization

High-throughput sequencing

General. For several decades, DNA was mainly sequenced by the method established in 1977 by Frederick Sanger (→56). With the start of the human genome project in 1990, but also due to the growing importance of DNA sequences in research and medicine, the needs for faster methods of sequencing grew. Today (2014), a genome of 10^9 bp can be sequenced in one day (the time needed to analyze the sequencing results is, of course, quite a different story!). The market volume of equipment and software for such high-throughput, "next generation sequencing" (HTS) was estimated at 1.5 billion US$ in 2013. Among a variety of methods that have been presented and realized, we will focus here on a) the 454 sequencer of Roche, b) the HiSeq analyzers of Illumina, c) the SOLID analyzer of Life Technologies, d) the RSII sequencer of Pacific Biosciences, and e) the nanopore method established by Oxford Nanopore Technologies.

454-Sequencer of Roche. By this method, originally developed by 454 Corporation in the USA, genomic DNA is sheared into small fragments to which terminal linkers are added. Through one of the two linkers, the fragment is then bound to micropearls on which each fragment is amplified by PCR. In the next step, each micropearl is deposited in a separate cavity of a picoliter plate, followed by pyroseqencing through primer extension. The read length is about 1 kb.

HiSeq analyzer of Illumina. Again, genomic DNA is sheared into fragments of 200–300 bp. Through ligation of appropriate adapters, bridged fragment clusters are obtained and deposited in a flow-through cell where they are amplified using solid-phase PCR. In the next step, double strands are formed in a highly parallel manner on each single-strand fragment using a specific fluorescence-labeled terminator for each base. As a consequence, the addition of each single base can be recognized by its specific fluorescence. Read length, however, is limited to about 200 b.

Solid Analyzer of Life Technologies Sheared fragments are separated into clonal fragments, ligated to the same P1 fragment, amplified by emulsion PCR and bound to a glass plate. Using a first primer for the P1 fragment, followed by n-1, n-2, n-3 and n-4 primers, a set of specific probes of the four bases each labeled with a fluorescent marker are competing for ligation to the next base and liberating their label once binding takes place. As this process is repeated several times with the n-x primers, accuracy is greatly enhanced. Read length is 75 to 100 bases.

PacBio RS II sequencer of Pacific Biosciences. Different to the hitherto described methods, in this procedure a single DNA-strand is sequenced. This strand passes a so-called SMRT cell (for "single-molecule real-time") in which DNA-polymerase is immobilized. Each of the four nucleotides required to form a DNA strand is modified with a different fluorescent marker and can freely enter he SMRT cell. Upon insertion of a labeled base, the marker is liberated, its fluorescence being propagated via a zero-modal waveguide. An amplification of the signal is not required. Since the core plate of the instrument contains one million SMRT cells, this procedure permits sequencing of large genomes. In addition, read length is over 5 kb.

GridION sequencer of Oxford Nanopore Technologies. In this system, single-stranded DNA is moved by electrophoresis, either through a biological pore (a recombinant porin in a double membrane) or through a chemical nanopore (graphene, silicium nitride). If a voltage is applied to the pore, the current changes during the passage of each nucleotide through the pore with a value that depends on geometry, allowing for the analysis each nucleotide in the base sequence. Even methylated bases, important for, e. g., the analysis of epigenetic modifications, can be identified by their geometry and thus by their current signals. The reading speed of a nanopore sequencer is ~300 bases per second.

Bioinformatics. An indispensable part of high-throughput sequencers are the bioinformatic programs that use information from sequenced overlapping strands to obtain information for genes, genomic and genome sequences (shotgun). The number of erroneous assignments of base sequences increases proportionally with the size of a genome. As a consequence, multiple sequence reads must be compared to exclude read errors. As repetitive sequences, which cause more reading errors, increase with the size of a genome, further measures for corrections are necessary.

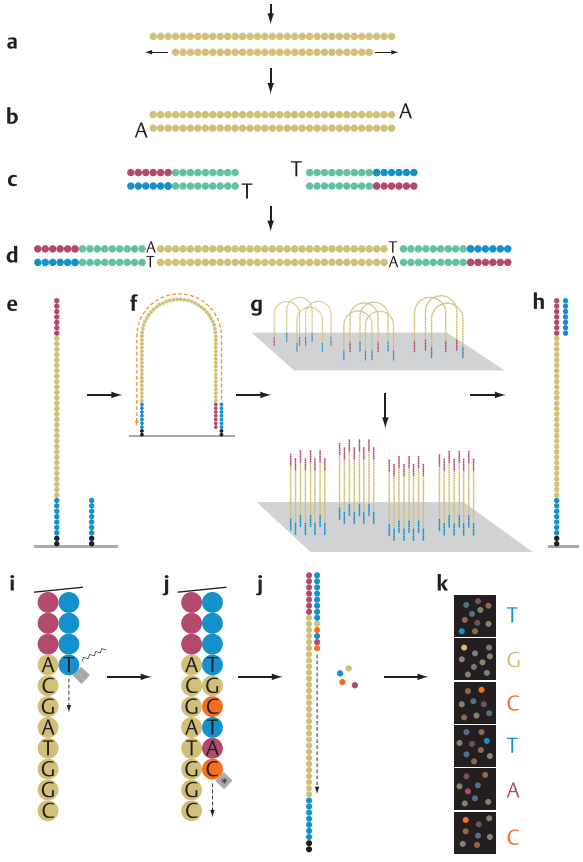
Highthroughput-Sequencing (HTS)

instrument	Roche 454 GS FLX	Illumina HiSeq	Life Tech Solid	Life Tech Ion Proton	Pacific Bio RSII	Oxford Nano-pore GridION
market introduction	2005	2007	2007	2013	2012	prototypes
price (2014) (US$)	500,000 (GS FLX+)	740,000 (HiSeq 2500)	665,000 (Solid 5500 xl W)	149,000	695,000	unknown
DNA preparation	F, L, A	F, L, amplifica-tion optional	F, L, A	F, L, A	F	F
read length (bases)	up to 1,000	up to 2 × 150 (sequencing from both ends)	1 × 75 or 2 × 50	200	5,500 – 8,500	very high
number of parallel reads	1 million	up to 8 billion	up to 2,4 billion	up tp 82 million	ca. 50,000	??
performance	700 million bases in 23 hrs	up to 1,000 billion bases in 6 days	240 billion bases in 5 days	10 billion bases in 4 hrs	375 million bases in 1 hr	300 bases/sec

The llumina HiSeq-X Ten system uses 10 instruments in parallel with 2 flow cells each (prize is 10 Mill. US-$, 2014). The cost for sequencing one human genome with this equipment is 1000 US-$

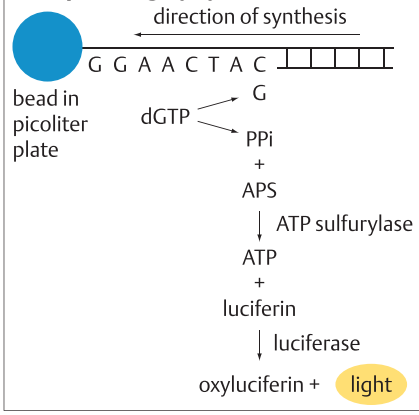
F = fragmentation, L = adapter ligation, A = amplification

Illumina sequencing method



*the 4 bases are ligated to different fluorescence markers

Sequencing by synthesis



library preparation: 6 h (3 h)
a fragment DNA
b repair ends/acid A overhang
c ligate adapters
d select ligated DNA

cluster generation: 5 h (3 h)
e attach DNA to flow cell
f perform bridge amplification
g generate clusters
h anneal sequencing primer

sequencing: 4–6 days (30 min)
i extend first base, read, and deblock
j repeat step above to extend strand
k generate base calls

duration/hands-on time

Proteomics

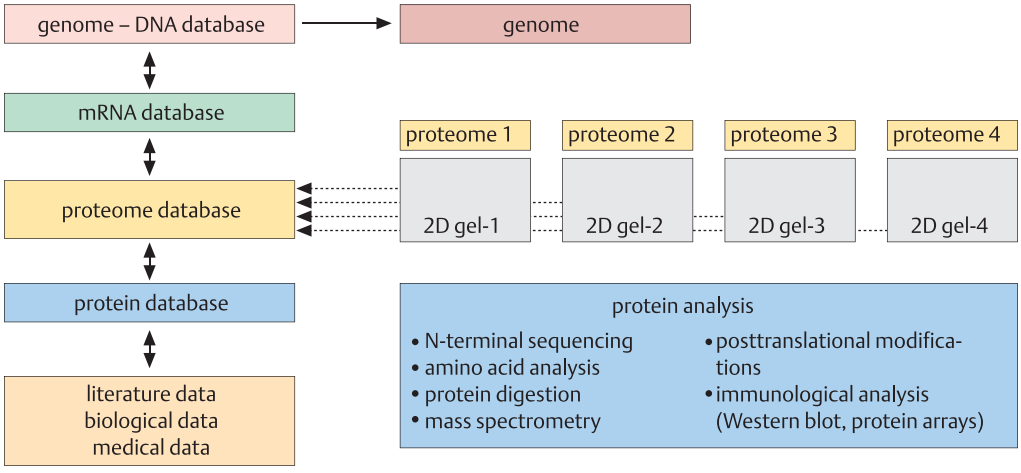
General. The term proteome was coined in 1995 and describes the total of all proteins encoded by a genome. In higher organisms, alternative splicing and post-translational modifications (→40) may lead, on average, to the formation of 10 proteins per gene. Proteomics is focused on understanding how proteins are differentially expressed and modified by the genome and how this determines their interaction and function (functional genomics).

Methods. The core procedure of proteomics research is the separation and identification of a large number of proteins. The *Escherichia coli* proteome, for example, contains ca. 4000 proteins. Sample preparation is critical and requires different protocol for membrane proteins than for cytoplasmic proteins. In eukaryotes, protein extracts of individual cell types are used preferentially, providing information on differential expression. The most important method for separating proteins is 2-dimen-sional polyacrylamide gel electrophoresis (2D PAGE), in which separation in the first dimension is based on isoelectric point, and in the second, on molar mass. Resolution can be improved by controlling the pH gradient. Semi-quantitative analysis of 2D PAGE gels is achieved by staining (silver), scanning, and computer analysis. High-throughput systems allow about 100 gels to be analyzed per week. Identification of rarely expressed proteins (10–1000 copies per cell), correlation of post-translationally modified proteins with the correct precursor protein, and quantification are major bottlenecks. One method of achieving quantitative analysis is to use recombinant antibody libraries, but the proteins' identities must be known. If quantities > 1 μg of protein are expressed, N-terminal sequencing may be possible if the N-terminus is not acylated; then the sequence can be matched with a computer database. For identification of less frequently expressed proteins, mass spectrometry is the method of choice. Using MALDI-TOF mass spectrometry (matrix-assisted laser-desorption-ionization time-of-flight), approximate molecular masses are obtained. Alternatively, proteins can be digested in the gel by trypsin, a protease that hydrolyzes peptides only at lysine or arginine residues, resulting in a peptide mixture whose mass again is analyzed by MALDI-TOF; comparison with a tryptic digest *in silico* provides information about the identity of the protein. This procedure can be used only if the genome sequence is available and all proteins have been assigned. If this is not true, electrospray mass spectrometry (ESI-TOF) can help to identify fragmentation patterns of the unknown protein, leading to protein sequence tags (PST), which can often be assigned to entries in a protein database. The sensitivity of both methods is several femtomoles per protein spot. To achieve this sensitivity, however, several hundred thousand copies of this particular protein must be expressed in the cell. Metal coded tagging (MeCAT) coding of proteins enhances sensitivity further to the attomolar (10^{-18} M) range; with this technique, peptides are labeled by lanthanides, and inductively coupled plasma mass spectrometry (ICP-MS) is used for quantification.

Applications. Using proteomics, cell-specific or induced variations in the expressed protein pattern of a cell can be analyzed. Examples are 1) differences in the expression pattern of the proteome of *E. coli* after growth on varied C source, 2) comparison of the protein patterns of pancreatic cells from a healthy and a diabetic person, and 3) toxicological investigations on altered protein patterns in the liver after medication or intoxication. Often, marker proteins for specific diseases have already been identified. Proteomics can also help to identify protein functions in a cell. To this end, scientists try to build a proteome map for the cell type under investigation and to understand protein interactions in this "cell factory". A useful method is the genetic introduction of polyhistidine tags, which can be used to purify the expressed protein and associated proteins by affinity purification and establish their identity by mass spectrometry.

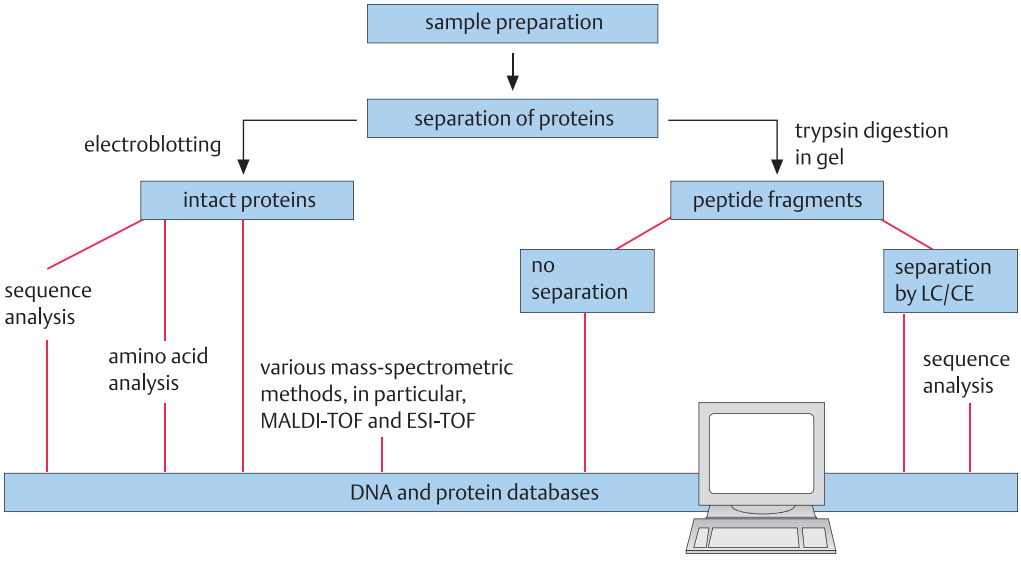
Human proteomics initiative. As of 2014, the Human Proteome Reference Database contained >30,000 annotated protein entries with about 94,000 post-translational modifications. These entries are citations, and 21,200 experimental or predicted post-translational modifications. Within the human proteome project, the proteins coded by each human chromosome are being elucidated, as are the proteins associated with tissues of clinical interest such as blood plasma, liver, brain and urine.

One genome – various proteomes

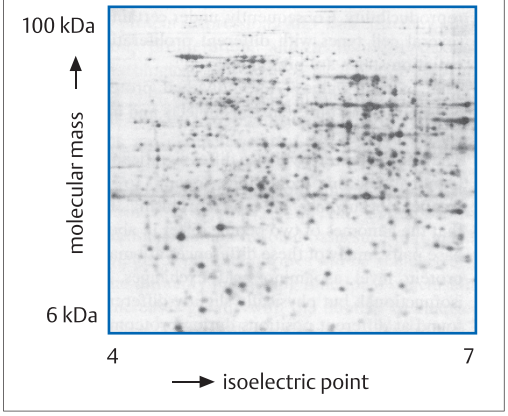


A higher organism with 100,000 genes may contain > 1 million protein species, which are expressed in modified forms by different cell types (> 50,000 protein species per cell type)

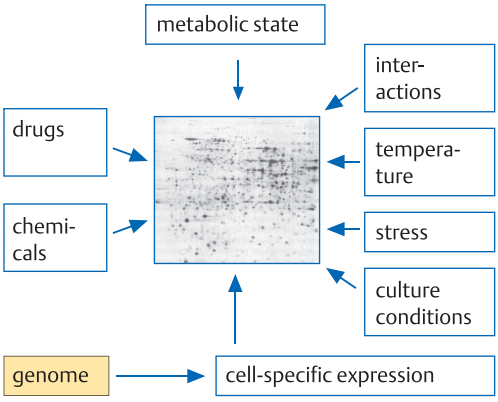
Methods of proteome analysis



2D gel of yeast proteins



Applications



DNA and protein arrays

General. DNA microarrays on solid surfaces (DNA chips, gene chips), or "liquid arrays", allow for simultaneous determination of many hybridization events. They are used for highly parallel genotyping, e. g., for detecting polymorphisms; for expression analysis, e. g., for studying differences in gene transcription; and for DNA sequencing. Commercially available DNA microarrays include "gene filters" made from nylon or nitrocellulose, which contain sets of cDNA fragments of yeast, mouse, man, or another organism. They offer densities of several hundred oligonucleotides cm^{-2}. Much higher densities of more than 10 000–100 000 oligonucleotides cm^{-2} are obtained on rigid surfaces such as glass or plastic by two alternative methods: 1) stepwise chemical synthesis of the oligonucleotides on the surface by photolithographic techniques ("*in situ* synthesis"); or by 2) microdeposition of DNA, RNA, or oligonucleotides with spotter equipment. "Liquid arrays," based on DNA-tagged microspheres, are a promising alternative for highly parallel DNA assays.

Oligonucleotide microarrays are prepared by coupling phosphoamidites, using nucleotide-specific protection groups that are removed by photochemical reactions. The distribution of each single step of synthesis over the whole array is controlled by photolithographic techniques (masks). This procedure enables the preparation of arrays with as many as 250,000 oligonucleotides cm^{-2}; recently, an array was presented containing 60 million DNA probes for human SNP analysis on a single wafer (20 cm diameter). The length of the oligonucleotides is usually limited to 25 bases or less because errors, due to incomplete coupling and deprotection, multiply during each step of synthesis. The manufacture of the many photolithographic masks required in this process results in high production costs for the prototype chip, but low costs for mass production.

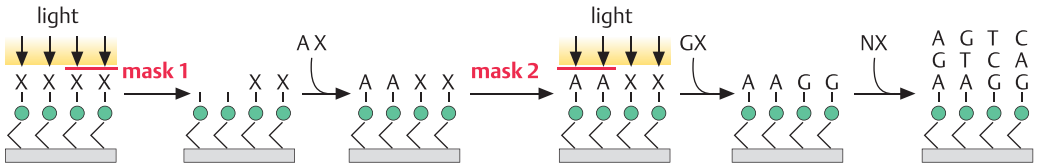
Microdeposition. Instead of synthesizing DNA *in situ*, single-stranded DNA or cDNA of any length, or synthetic oligonucleotides, can also be microdeposited on a surface such as glass. For the coupling step, standard surface-chemistry procedures can be used. The polynucleotide is deposited either by contact spotting with pins or by non-contact spotting with drops, based on piezoelectric methods (inkjet printer technology). Commercial microspotters can reach speeds of >10,000 DNA or cDNA fragments h^{-1}. Such methods have allowed arrays with ca. 10,000 oligonucleotides cm^{-2} to be prepared. Oligonucleotide arrays are useful for detecting incomplete hybridization due to mismatches (e. g., SNP analysis). Another application is the resequencing of DNA. To this end, the array must contain a complete single-stranded sequence of the target DNA in the form of overlapping fragments. This approach has been validated by sequencing human mitochondrial DNA (mtDNA), using an array of 16,000 20mer oligonucleotides: 179 of the 180 known polymorphisms of mtDNA could be demonstrated in a single hybridization experiment.

Detection. Hybridization events are detected by labeling with radioactive, fluorogenic reporter groups, usually incorporated into the DNA probe during amplification. Detection is carried out by autoradiography, laser scanners, or CCD image analyzers. Biotinylated nucleotides can also be applied, which react, after hybridization, with gold-labeled avidin. The sensitivity of this method can be 50-fold increased if colloidal silver is used in the final step. For label-free detection of hybridization events, the masses of primer-specific extension products of the probe can be separated and analyzed using laser-desorption mass spectrometry (MALDI-TOF).

"Liquid arrays." Single-stranded DNA-tagged polystyrene microspheres have been differentially labeled with two fluorophores, allowing the rapid sorting of up to 100 different DNA sequences by their spectral addresses in a fluorescence-activated cell sorter (FACS).

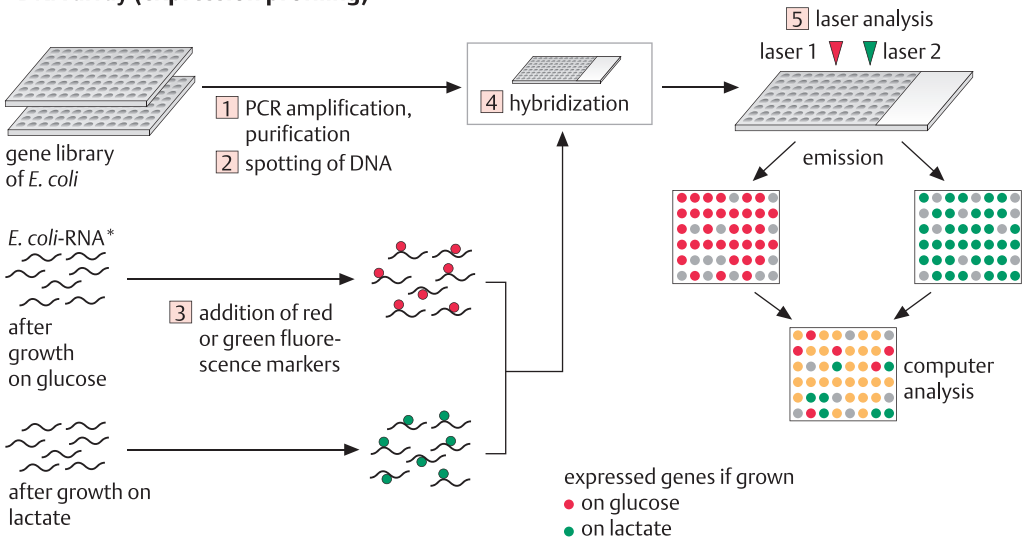
Protein arrays. Instead of DNA, a large number of proteins can be immobilized on a glass surface. For example, target proteins of kinases can be tested if all putative targets are immobilized on one array and kinase-catalyzed phosphorylation is carried out in the presence of ^{32}P-labeled ATP, followed by autoradiography. Libraries of recombinant antibodies (→244) or aptamers (→42) are often used for detection. The main applications of protein arrays are in blood serum diagnostics, the analysis of protein patterns in cell lysates, and in the analysis of protein-protein interactions.

Oligonucleotide arrays



mask 1, 2: photolithographic masks
AX, GX, NX: phosphoamidites of **a**denosine, **g**uanosine, any **n**ucleotide

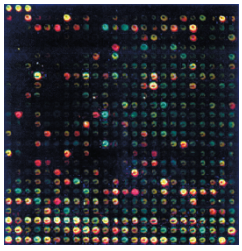
DNA array (expression profiling)



*RNA is retrotranscribed into single-stranded DNA

Applications and detection methods

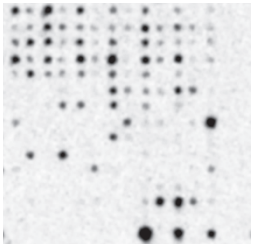
array density	applications
high, > 10 000 DNA sequences	identification of genes, transcription (expression) profiling
medium, 1 000 – 10 000 DNA sequences	analysis of mutations, alleles, polymorphisms
low, < 1 000 DNA sequences	genetic predispositions, acute diseases, infections



DNA microarray

hybridization with fluorescence- or gold-/silver-labeled DNA

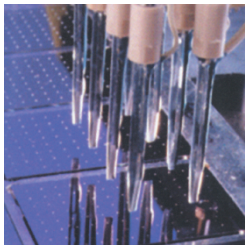
detection by CCD, laser fluorescence scanning or absorption (silver)



DNA filter assay

hybridization with radioactively labelled DNA

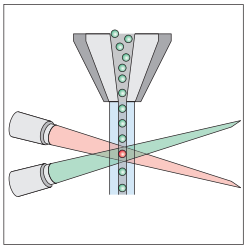
detection by autoradiography



mass-spectrometric DNA assay

single-stranded DNA is probed by primer-specific extension

detection by sequence-specific molecular mass using MALDI-TOF mass spectrometry



spectral DNA assay ("liquid array")

microspheres are labelled with single-stranded DNA probes and coded with 2 fluorophores

detection by interrogating spectral addresses in a FACS system

Metabolic engineering

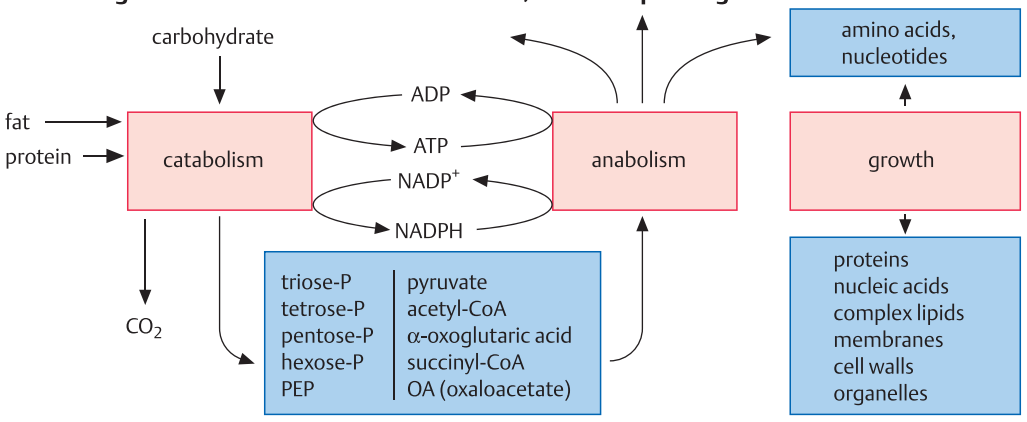
General. Increasingly in-depth knowledge of the metabolism, its pathways, and their regulation has led to attempts to describe the entire metabolic flux of whole cells, using mathematical modeling, computer simulations, and fast, precise measurement techniques. Based on these analyses, investigators are attempting to modulate metabolism, regulatory functions, and signal networks, to direct the metabolic flow towards a desired product or to adapt it to a desired carbon source (metabolic engineering). The experimental setup usually includes genetic engineering of a key step in synthesis or catabolism (metabolic design).

Metabolic flux analysis. Once the biochemistry of a metabolic pathway (→36) is understood, a system of balance equations can be formulated for each intermediate and added up in a stoichiometric matrix. Using these equations under a given set of conditions, the metabolic flux within the pathway can be calculated. A necessary condition is that mass transfer to and from the environment through the cell membrane (substrate uptake and product release) must be quantified. This may not be sufficient, however, to solve the system of balance equations unambiguously. In the standard case of metabolic fluxes through additional pathways, it is necessary to determine enzyme activities, to measure expression profiles of genes using DNA arrays (→316), or to carry out marker experiments using probes that are labeled with radioactive (^{14}C, ^{3}H, ^{32}P, or ^{35}S) or, preferably, stable isotopes (^{13}C or N^{15}). To this end, cells are grown on well defined, isotope-labeled substrate mixtures, and the metabolites, occasionally also the macromolecules, are analyzed for their isotope distribution, using nuclear magnetic resonance (NMR) or mass spectrometry (MS). In favorable cases, the use of such isotopomers (isotope-labeled isomers) may allow generation of a mathematical model for metabolic flow in a branched pathway or even in a whole cell.

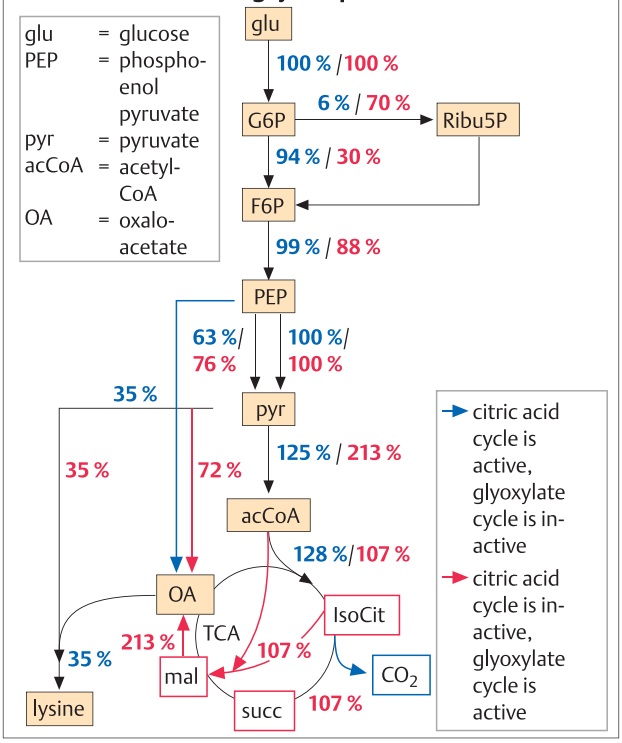
Metabolic control analysis. If one wants to understand which enzyme in a pathway is limiting to the metabolic flux, the hierarchy of flux control must be analyzed. In other words, we need to understand which fractional change in the activities of all the enzymes involved has the largest impact on the overall pathway. Usually, this type of control is distributed among several enzymes. To determine each enzyme's control coefficient, its expression can be modified by genetic engineering, and the resulting change in metabolic flux can be quantitatively analyzed. Alternatively, the flux control coefficients can be calculated on the basis of mathematical models that adequately describe the kinetic parameters of the pathway network under investigation. For experimental validation of such models, intracellular metabolites are measured under transient process conditions. This is done, for instance, by stimulating a stationary cell culture by adding a substrate (e. g., a glucose spike) and measuring the response of the cell by analyzing all relevant internal pool concentrations.

Applications. In microorganisms, these methods are mostly used to 1) enhance the spectrum of substrates, 2) enlarge their biodegradation potential, and 3) increase yields for the production of metabolites. For instance, the lactose operon was successfully cloned and expressed in *Zymomonas mobilis*, allowing for the use of whey for ethanol production (→138), and in *Corynebacterium glutamicum*, for glutamic acid (→126) and lysine (→128) production from whey. By cloning genes for the catabolism of aromatic compounds into *Pseudomonas* sp. B13, engineered mutants were obtained which, unlike the wild strain, could grow on chlorinated or methylated aromatic compounds. The increase in product formation by metabolic engineering is studied for many purposes, e. g., for the production of amino acids, ethanol, biopolymers, and vitamins, and also for producing secondary metabolites such as antibiotics. For example, a thorough analysis of competitive metabolic fluxes during the biosynthesis and secretion of L-lysine in *Corynebacterium glutamicum* led to a strategy that increased lysine yields by 50 %, by using molecular genetics for redirection of flux. An essential aspect of metabolic engineering is determining the optimal genetic manipulation when altering fluxes. Computational algorithms have been developed that provide recommendations for the overexpression, knock-out or insertion of specific genes involved in the pathway being considered. Companies such as Insilico Biotechnology provide simulation platforms for the modeling of dozens of microbial and animal cells in view of optimizing their metabolism.

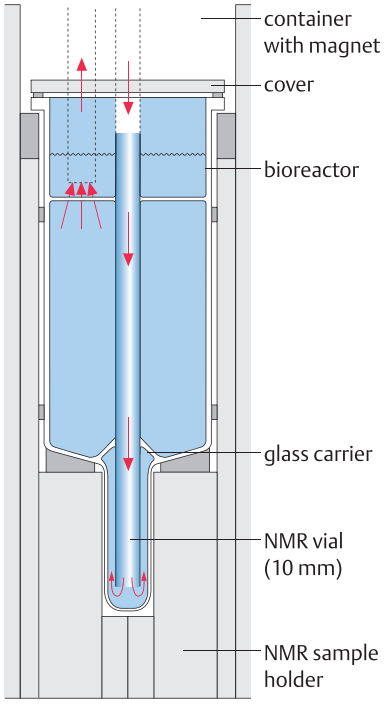
Block diagram of the metabolism of an aerobic, heterotrophic organism



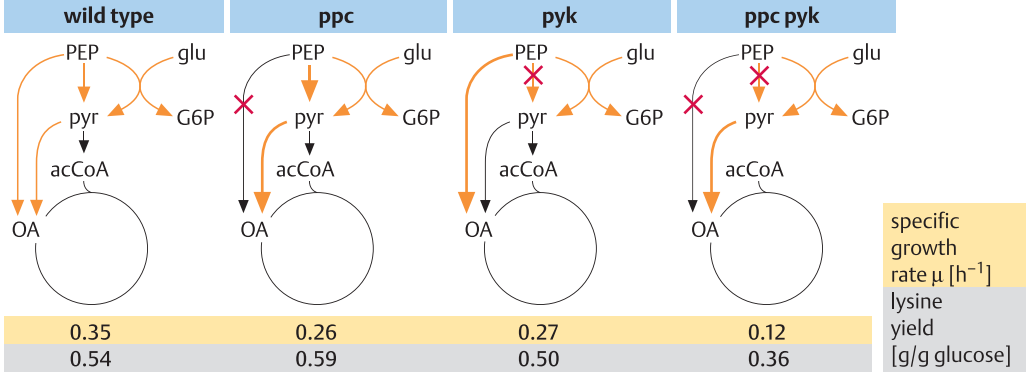
Metabolic flux during lysine production



Metabolite analysis using NMR



Analysis of metabolites



Synthetic biology

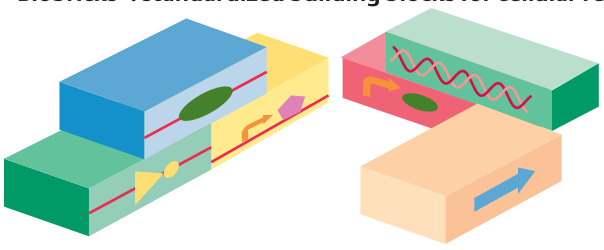
General. Several methods are summarized under this term, such as the use of chemically modified bases in triplet codons, which enlarge the genetic code, or the design of artificial enzymes using bioinformatics. The most frequently used concept for a synthetic biology, however, is the construction of artificial biological circuits on plasmid DNA ("Biobricks") that are then cloned into a host organism (usually *E. coli*), following the concept of a Lego model. In view of technical applications, two variations of synthetic biology are most important: a) the construction of host organism with a synthetic or a reduced genome, and b) the construction of artificial metabolic pathways in microorganisms.

Host organisms with a reduced genome. An early example for such work is the construction of the attenuated strain *E. coli*-K12 (→20) where all plasmids and all putative genes coding for gene products related to pathogenicity in humans such as fimbrial adhesins, surface antigens, etc. have been deleted, thus creating a safe strain for genetic engineering. In recent years, microorganisms with more extensive deletions in their genomes were created in view of improved product formation. Thus, a strain of *Corynebacterium glutamicum* (→20), a bacterium which is used for the commercial production of glutamic acid, lysine and other products, has been subjected to extensive genome reduction and turned into a universal host for enzymes (e. g., transglutaminase) or therapeutic proteins (e. g., hEGF) that are correctly folded and exported into the surrounding medium (CORYNEX™). In the case of *Schizosaccharomyces pombe* (→14), many of the 59 proteases as well as other proteins coded by the genome, were removed, resulting in a mutant organism that is able to secrete correctly folded and highly pure enzymes or therapeutic proteins (e. g., human growth hormone, transferrin) into the medium (ASPEX™). The genome of a *Bacillus subtilis*-strain (→20) was reduced by 21 % and modified in areas important for substrate uptake and product secretion. As a result, secretion of a foreign cellulase that is used as a detergent enzyme, was increased by 250 %. Other achievements that are closer to basic research concern the construction of "artificial" or chimeric microorganisms. Thus, in 2010, a *Mycoplasma mycoides JCVI-syn1.0* was described, whose fully-synthetic 1.08 Mbp genome was transferred into an enucleated *M. capricolum* strain: the "artificial" bacterium grew and divided. A vital bacterial chimera was created in 2012 by inserting the whole genome of the cyanobacterium Synechocystis PCC6803 in *Bacillus subtilis* strain 168.

Construction of new metabolic path-ways. An early industrial example (2005) is the construction of a synthetic pathway for the synthesis of 1,3-propanediol (→142) in *E. coli*. To achieve overproduction of this industrial chemical, used for polyester synthesis, two genes from baker's yeast and one from *Klebsiella pneumoniae* were introduced, and the metabolism was optimized by metabolic engineering. A new example (2014) is the industrial production of hydrocortisone (→252) by baker's yeast, starting from ergosterol, a cell-wall component of yeast. To facilitate this pathway, several genes of beef, human and plant origin were introduced, and several steps involving the transfer of intermediates into compartments such as the mitochondria had to be optimized, which took about 15 years to achieve. Another example is the synthesis of artemisinin. This anti-malaria drug is formed by the annual wormwood plant (*Artemisia annua*) but is difficult to obtain from the plant or by plant tissue culture. New synthetic pathways were constructed both for *E. coli* and for baker's yeast, using a gene from the green algae *Haematococcus pluvialis*, which led to good yields of the precursor compound amorphadiene in these microorganisms. Some cyanobacteria produce high levels of triglycerides, useful for the manufacture of fatty acid methyl esters ("biodiesel"). If their triglyceride pathway is complemented by a fatty acid-CoA-reductase and a fatty aldehyde decarbonylase, microbial alkanes are formed (→18).

Risks and ethical concerns. (→332) Using methods of synthetic biology, viruses pathogenic to humans were also synthesized (polio, Spanish flu), making it obvious that rules and protective measures are necessary to prevent misuses of this new technology. Ethical concerns arise also from the potential to modify stem cells. Ecological and IP issues relate to seed, industrial microorganisms and the deliberate release of "artificial" microorganisms into the environment, which is under a moratorium.

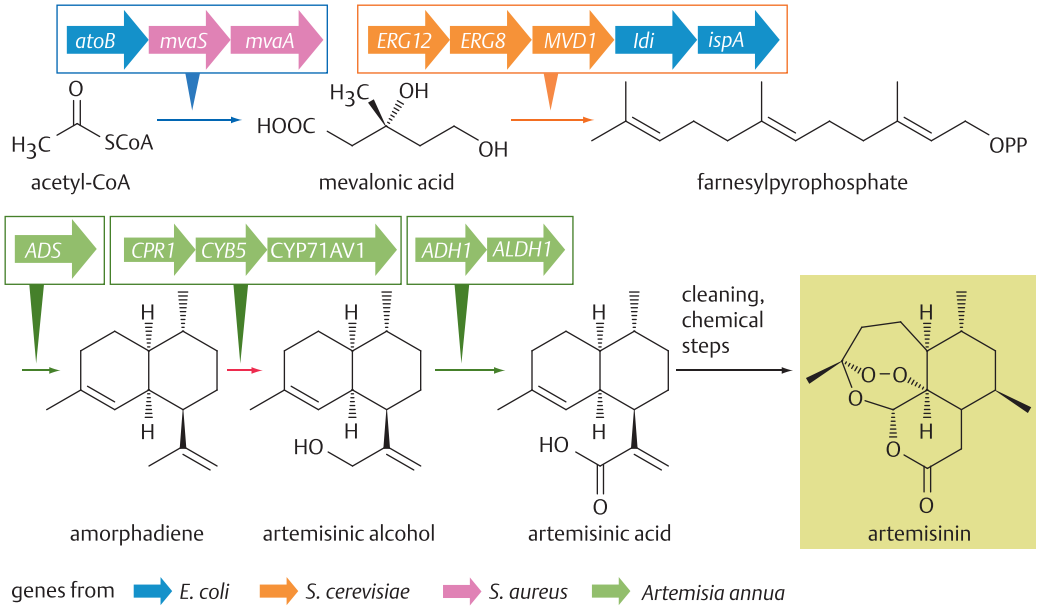
"Biobricks": standardized building blocks for cellular reactions



DNA coding for a BioBrick part is inserted into a plasmid, following specified rules, and can be recombined in the host organism with other bricks

symbol	"Biobrick"
	promoter
	coding sequence
	RBS
	inverter
	DNA
	terminator

Synthese of artemisinin, an antimalaria drug

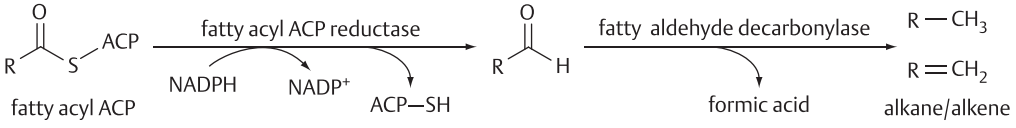


For manufacturing, *S. cerevisiae* is used as host organism. Yields of artemisinic acid are ~ 25 g/L

Examples for "synthetic" metabolic pathways

product	host organism	additional genes from	research group
artemisinin	*S. cerevisiae*	bacteria, plants	Jay Keasling, U California
alkenes	Cyanobacteria	plants, algae	Andreas Schirmer, LS9 Inc
1,3-propanediol	*E. coli*	*S. cerevisiae*, Klebsiella	C E Nakamura, Dupont
3-hydroxypropionic acid	*E. coli*	Klebsiella, Azospirillium	David H. Cameron, Cargill
hydrocortison	*S. cerevisiae*	bovine, plants	Denis Pompon, CNRS
ascorbic acid	*C. glutamicum*	*Erwinia herbicola*	Genentech, 1983
indigo	*E. coli*	*Pseudomonas putida*	David Gibson, 1983

Alkanes/alkenes form fatty acids



Systems biology

General. Systems biology is a new field of research aiming at a holistic description of cell functions. It is based on the functional analysis of metabolic, regulatory, and signaling compounds, which are aggregated to functional modules. These functional modules are combined to provide a cell model that is based on experiment and allows interactive predictions of a cell's behavior. A long-term objective of systems biology is, for example, the support of clinical trials by simulation studies. Using an analogy, systems biology is a dynamic roadmap of the cell, which includes traffic patterns, why such patterns emerge, and how we can control them.

Key components. Before compiling information on a biological system, detailed data must become available on the *system's structure*. This includes the functions and interactions of genes, protein structures, and biochemical pathways and the mechanisms by which the intracellular and multicellular structures are modulated. A second key property of living cells is their capacity to undergo dynamic changes, in response to internal or external factors. *Simulation studies* describe these changes and allow the identification of essential mechanisms underlying specific behaviors. *Control factors* that minimize malfunctions must be elucidated. Finally, *design principles* and simulations help to cluster the above modules into a functional system.

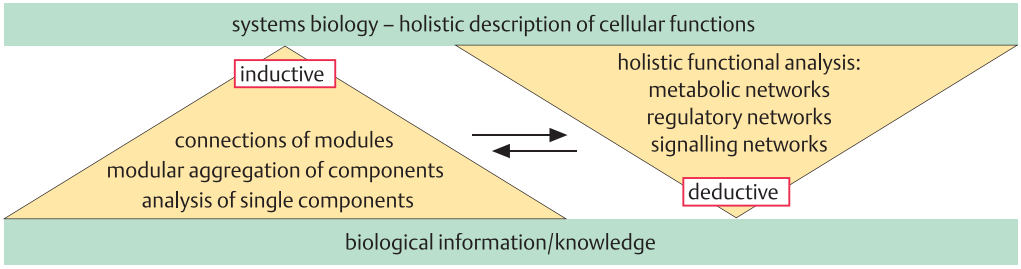
Measurement. Any analysis of the biological regulation present in a complex living system requires extensive databases, such as the sequencing and expression profiling of genes (genome and transcriptome) (→316), proteome analysis (→314), and the simultaneous measurement of enzyme activities and metabolites (metabolome) (→318) within milliseconds, to follow metabolic events under defined conditions in a highly parallel format with high throughput (metabolomics). Single-molecule measurements, robotic nano-devices, and femtolasers that permit visualization of molecular interactions are typical examples ("interactome"). Microfluidic systems such as the micro-Total Analysis System (μ-TAS) allow picoliters of samples to be measured rapidly and precisely, and many established techniques such as PCR and protein separation by capillary electrophoresis are being scaled-down to follow, for example, mRNA levels and protein modifications.

Robustness. Biological systems exhibit a considerable robustness, which protects the cell if a subsystem runs out of control. Phenomenological parameters related to robustness are: 1) the ability to cope with environmental changes, 2) a system's relative insensitivity to kinetic changes, and 3) a graceful degradation of the system's functions after damage, rather than catastrophic failure. This behavior is attained by various control systems such as negative feedback and feedforward control, by redundancy and structural stability of vital functions, and by modularity, that is, physical or functional insulation of subsystems so that failure of one module does not spread to others. Related systems are also used in the engineering of complex machines.

Computational tools originally designed for general engineering purposes are frequently used in systems biology. Because data integration, management, visualization, and analysis of large-scale cellular systems is a challenging task, integrated modeling and simulation software has been developed. More recently, Systems Biology Mark-Up Language (SBML), along with Cell Mark-Up Language (CellML), have evolved as promising standards for XML-based computer-readable model definitions. These standards enable models to be exchanged irrespective of the computational tool used. Systems Biology Workbench (SBW) is built on SBML and provides a framework of modular open-source software that can be shared among research institutions for collective projects.

Consortia and applications. Major projects in systems biology are under way (2014) in many countries, with a strong focus on health-related issues. In the USA, the Institute of Systems Biology in Seattle, as just one example, does research on perturbations of biological networks that result in various diseases. In Germany, major projects were funded for a systems understanding of the biological network of the liver, and on the perturbations of networks in cancer cells or on aging. In Japan, the Systems Biology Institute covers medical projects as well, but also reaches out to understand the systems biology of microorganisms such as *E. coli* or yeast, or even of whole consortia such as coral reefs.

Systems biology – the basic concept



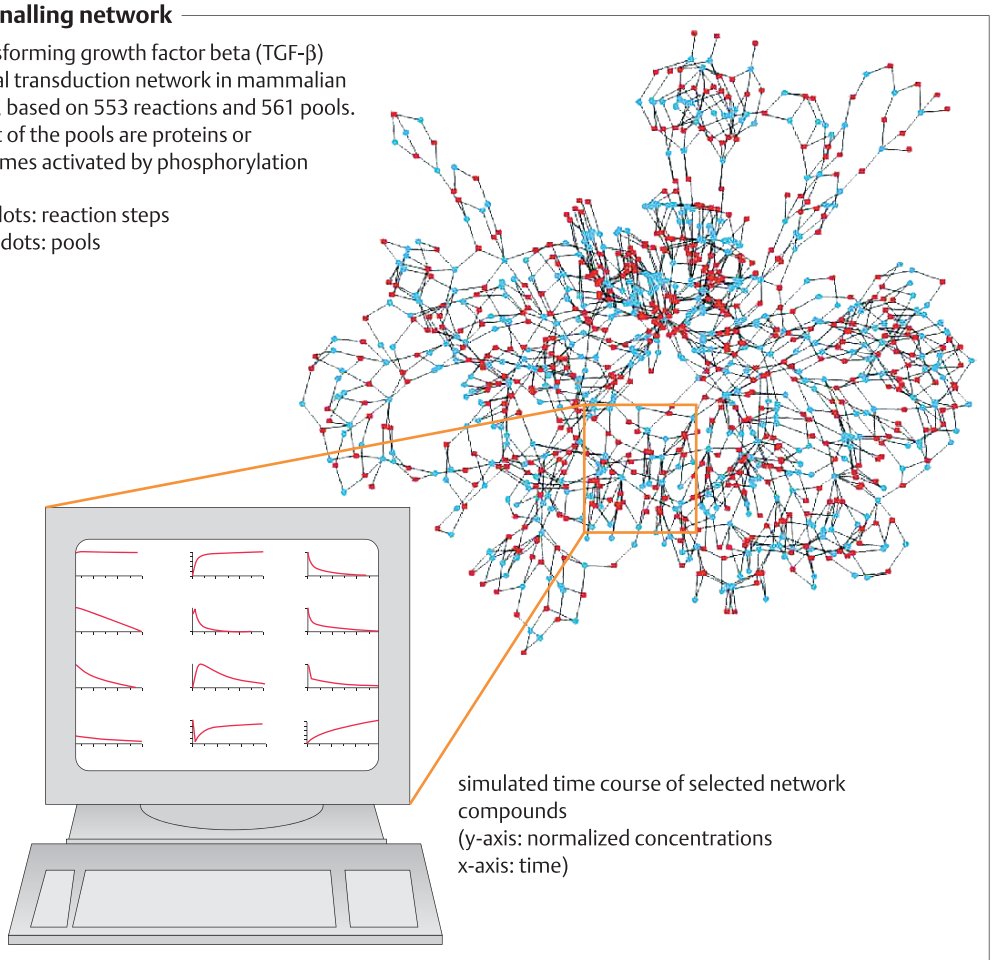
Terms and experimental methods

term	quantitative data on	methods
metabolome	metabolites	robot-assisted rapid LC – MS, enzyme assays ^{15}N ^{13}C tracer analysis based on NMR
genome	gene function	knock-out experiments
transcriptome	differential mRNA formation	expression profiles using cDNA arrays
proteome	differential protein synthesis and modification	MALDI-TOF or electron spray mass spectrometry, 2D gel electrophoresis
"interactome"	differential protein-protein interactions	two hybrid system, atom force microscopy, fluorescence resonance electron transfer (FRET)

Signalling network

transforming growth factor beta (TGF-β) signal transduction network in mammalian cells, based on 553 reactions and 561 pools. Most of the pools are proteins or enzymes activated by phosphorylation

red dots: reaction steps
blue dots: pools



simulated time course of selected network compounds
(y-axis: normalized concentrations
x-axis: time)

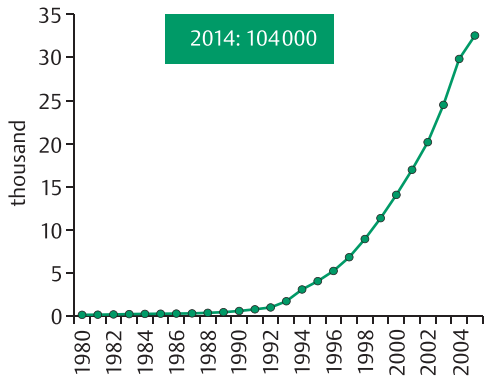
Bioinformatics: sequence and structural databases

General. The rapid progress of molecular biology would be unthinkable without the breathtaking developments in computing and telecommunications – two technologies that ensure storage, rapid sorting, and worldwide retrieval of large amounts of biological data. The communication platform for data exchange is the Internet, founded as early as 1980 for scientific telecommunications within the USA. Today, the Web spans the globe with >25 billion indexed websites (2009) and >2 billion users (2010), a figure that is still increasing every year. Apart from its commercial uses, the Internet enables the global exchange of scientific data and their collective or individual retrieval and organization. The terms bioinformatics and biocomputing encompass many aspects, including the handling and processing of information on DNA and protein sequences, protein structures, and metabolic pathways.

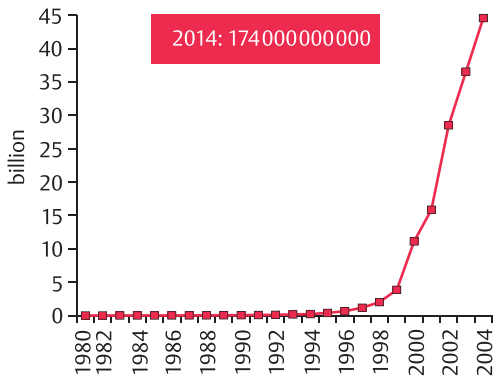
Sequence information. The broad implementation and high speed of DNA sequencing has led to a steep increase in the number of stored DNA sequences available to the public, about 10-fold every two years. By the end of 2004, the total number of DNA nucleotide bases that were sequenced exceeded 45 billion; with the arrival of high-throughput sequencing (→312), they rose to 10^{11} bases (100 Tb) in 2011. The storage of these data in a single global databank ("GenBank"), with three mirror sites in the USA, Japan, and Europe (UK), and their access through the Web, permits scientists to compare new DNA sequence data in real time with known sequences and to use this information to reach conclusions regarding any identity or homology between their DNA or protein sequences and previously defined genes and proteins. BLAST (→326) (basic local alignment search tool) is a type of algorithm that identifies the homology of an experimentally determined sequence to similar sequences stored in GenBank, providing information on putative functions. Protein sequence information is stored by the UniProt consortium, which was formed from the SwissProt database in Geneva, the European Bioinformatics Institute in Hinxton, UK, and the Protein Information Resource (PIR) in Washington, D. C. It contained >500,000 annotated sequences in early 2014.

Structure information. Relative to sequence information, data on tertiary structures of proteins have grown less quickly, because they rely essentially on x-ray analysis, which requires time-consuming preparation of crystals and their heavy-metal derivatives. This experimental bottleneck is a particular drawback for the analysis of large protein complexes and for membrane proteins, which constitute about 30 % of all known proteins. Protein structure analysis by multidimensional NMR, an alternative procedure applicable to proteins in solution, is presently limited to proteins with molecular mass <30kDa. In spite of these experimental difficulties, the number of protein structures available from the Protein Data Bank (PDB, Research Collaboratory for Structural Bioinformatics RCSB) by early 2014 came close to 100,000 protein structures and structure variants, with an increase of about 8000 per year. If the homology of a protein sequence (for a protein of unknown tertiary structure) to the sequence of a protein of experimentally determined structure exceeds 30 %, a structural model of the unknown protein can usually be built on the computer, using homology modeling. In other structure databases (SCOP, CATH) one can find, at present, about 100 to 150 different protein architectures. Although the number of theoretically possible protein sequences is astronomically high (for a protein made up of 300 amino acids, the possible number of different sequences is 20^{300}), two secondary structural motifs (the α helix and the β sheet) (→28) are relatively stable. In addition, larger structural motifs seem to be constituted in a modular manner. This greatly reduces the number of probable protein architectures and allows the comparison of such motifs with a dictionary of known structural elements in proteins (DSSP). An empirical method that has been widely and successfully applied is the generation of structure-function databases of a particular type of protein. As an example, the cytochrome P450 (CYP) Engineering Database (CYPED) contains over 13,000 CYP sequences and 741 annotated structures (2014), which provide valuable information on the structural features of newly sequenced CYPs and rational starting points for their protein engineering (→198).

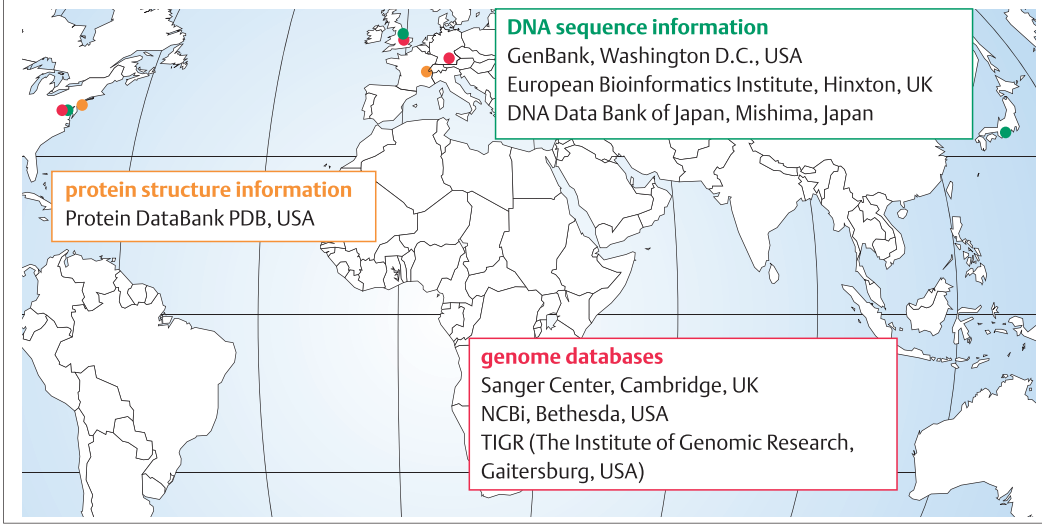
Number of Internet hosts



Number of bases in GenBank



Important hosts for bioinformatics



Important websites (a small selection)

DNA sequences		
GenBank European Nucleotide Archive (ENA) DNA-Databank of Japan (3 *mirror sites*)	~ 174 billion base pairs (August 2014)	www.ncbi.nlm.nih.gov/genbank.html www.ebi.ac.uk/ena.html www.ddbj.nig.ac.jp
genome informations		
Sanger Welcome Trust Institute	human genome, mouse genome, many other genomes	www.sanger.ac.uk
National Center for Biotechnology Information (NCBI)	human genome, mouse genome, many other genomes	www.ncbi.nlm.nih.gov
protein structures		
RCSB Protein Data Bank	~ 104,000 protein structures (October 2014)	www.rcsb.org/pdb/home/home.do
Uni ProtKB	~ 500,000 protein structures and sequences (Oct. 2014), ~ 140,000 annotated human protein sequences	www.uniprot.org
CYPED CYP Engineering Database	CYP structure and function: >13,000 P450 (CYP) sequences >700 functionally annotated structures	www.cyped.uni-stuttgart.de

Bioinformatics: functional analyses

General. Apart from algorithms or databases on gene sequences or protein sequences and structures, there are many useful bioinformatic tools concerning biochemical functions. This short review will focus on a) methods for genome annotation, b) the metabolic database KEGG, c) the FANTOM consortium which analyzes transcription-related functional elements (→62) d) the enzyme database BRENDA, and e) genome-wide association studies.

Genome annotations assign functions to DNA sequences in a genome. They usually comprise three steps: 1) elimination of those DNA sequences that do not code for proteins, 2) the prediction of functional genes, and 3) the association of biological function to such functional genes. For the prediction of functional genes, open reading frames (ORFs) that contain regulatory sequence motifs but no stop codons are characteristic. Annotations begin with computer analysis and are then complemented by manual annotation (curation), which involves human expertise. A key tool for the prediction of biological functions is **BLAST** (**B**asic **L**ocal **A**lignment **S**earch **T**ool), an algorithm that compares a new DNA sequence with sequences contained in databases (usually GenBank) and proposes putative gene functions using a ranked list of sequence homologies. A gene ontology database allows the cataloguing of diverse biological gene functions within a universally valid database that is independent of the organism in which this function was first identified. This also includes biological functions of the gene product. Over 98 % of the human genome – and of the genomes of most higher organisms – does not code for proteins. The ENCODE consortium attempts to determine the role of this "exome."

KEGG (**K**yoto **E**ncyclopedia of **G**enes and **G**enomes) attempts to describe whole cells and organisms using bioinformatic tools. This comprehensive encyclopedia contains a host of information on structure, reaction equations, metabolic pathways and functional hierarchies. In addition, information about human diseases and drugs for their treatment are stored. Using the *KEGG Pathway Database*, metabolic pathways of an organism can be derived from its genome sequence, and displayed as a metabolic network (→36). Metabolic pathways of different organisms can be automatically compared using the *KEGG Orthology (KO)* system. Furthermore, cellular processes, e. g., the transport among different compartments or organismal systems such as the T-cell receptor cascade or the insulin signal cascade, can be extracted from a genome sequence, and graphically presented. In the *KEGG Disease Database*, numerous human diseases are presented on the basis of gene defects or disturbed metabolic networks, often in combination with the targets of drug interventions.

FANTOM (**f**unctional **an**notation of the mammalian gen**om**e) is an international consortial project coordinated out of Japan. It uses a CAGE algorithm (**c**ap **a**nalysis of **g**enetic **e**lements) to analyze transcription factors (→62), promoters and other genetic elements that take part in transcription, and is applied to cells from healthy and sick individuals, both from humans and mice. Recently (2014), a comprehensive mapping of transcription factors of all human cell types was published by the FANTOM consortium.

BRENDA (**B**raunschweig **En**zyme **Da**tabase) is an online information system about enzymes and metabolic pathways. It contains about 6,500 enzyme types, classified by EC-numbers, and their properties as obtained by text mining from about 130,000 publications in PubMed. Ligands of all enzymes (>1.4 million enzyme ligands such as substrates, co-substrates, cofactors, inhibitors etc.) can also be searched using BRENDA.

Genome-wide association studies serve to link a specific phenotype – usually a multigenic disease – to related genetic markers, e. g. SNPs. In some cases, the resolution is already good enough to describe a human disease in terms of polymorphisms of genetic elements such as functional genes, promoters etc (→298). This progress, and the quickly decreasing cost for high-throughput genome sequencing (HTS), have led to first commercial offers for an analysis of genetic disposition of diseases (e. g., 23andMe), which are, however, not authorized by FDA or EMEA (2013). Special procedures, e. g., on SNPs in mitochondrial DNA, are already used in molecular genealogy and anthropology, e. g., in the human genographic project supported by the National Geographic Society and IBM.

BLAST analysis of a gene sequence

ATGGCATTGCCGTCAGGTTCTGACCCGGCCTTTAGCCAGCCGAAG
unknown sequence obtained from DNA sequencing

translated into protein sequence

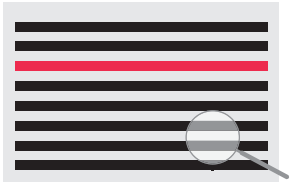
MALPSGSDPAFSQPKSVLDAGLTCQGAS

"nucleotide blast"

"blastx"

comparison with database 1
(180 M nucleotide sequences*)

comparison with database 2
(52 M proteine sequences)*

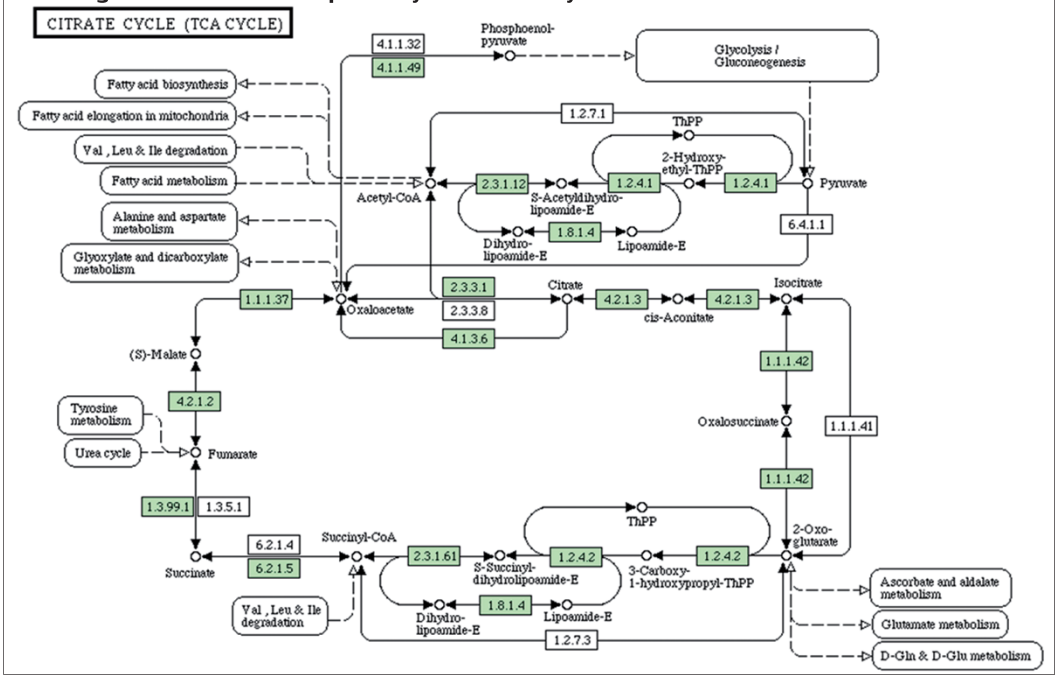


*as of 8/2014

hit	"score"	E-value	identity	sequence
Candida antartica lipase B	1,775	0.0	100%	ATGGCATTGCCGTCAGGTTCTGAC

hit	"score"	E-value	identity	sequence
Candida antartica lipase B	607	0,0	100%	KSVLDAGLTCQGASPSSVSKPILLV
putative lipase	488	4e-170	77%	KSVLDAGLTCQGTSPTSVTKPILLV
putative lipase	414	6e-141	67%	KSLLDAGLTCQNGSPSSQSKPILLV
lipase	145	1e-36	40%	KAQLDSVLACQNGSPSSQKNPILLV

KEGG generated metabolic pathway for the TCA cycle in *E. coli*



BRENDA: different enzymes involved in alcohol oxidation

EC number	enzyme name	reaction
1.1.1.1	alcohol dehydrogenase (NAD^+)	primary alcohol + NAD^+ = acetyldehyde + NADH + H^+
1.1.5.5	alcohol dehydrogenase (quinone)	ethanol + ubiquinone = acetaldehyde + ubiquinol
1.1.99.36	alcohol dehydrogenase (nicotinoprotein)	ethanol + acceptor = acetaldehyde + reduced acceptor

Carbon sources (C-sources)

General. The predictable shortage of fossil raw materials has stimulated global research activities on the use of biological feedstocks or CO_2 for the production of biofuels and chemical raw materials. Currently, starch and saccharose (→176) are used for the industrial manufacture of fermentation products such as bioethanol (→138) or dicarboxylic acids. Fats and oils, a different kind of sustainable raw material, are used to produce fatty acid methyl esters ("biodiesel") (→162) and "oleochemicals." The dual use of these raw materials for nutrition and in chemical technology has led to a global discussion about an "eat or drive" antagonism, in view of securing the food supply for a still-rising world population. Against this background, present discussions (2014) focus on 1) the use of CO_2 for the production of biomass and chemicals, using autotrophic microorganisms or alga (→18), 2) the pulping of biomass to fermentable sugars, followed by fermentation processes, 3) the use of lignin as a raw material, and 4) the industrial use of waste fats and oils which are not used in food production, e. g., gutter oil.

CO2 as a C-source. The largest part of all biomass is formed by plants or algae using CO_2 and solar energy. Algae are presently explored (2014) (→18) as sources of liquid biofuels, e. g., of squalene or farnesene, but also of alkanes derived from triglycerides *via* synthetic biology. Using genetic engineering, the metabolism of higher plants can be rerouted as well. Thus, Eucalyptus trees or poplars were engineered and bred to grow faster, contain more polysaccharides and less lignin and are more suitable both for the production of pulp (→184) and for the sustainable production of biomass to be used as a carbon or energy source.

Synthesis gas (syngas) as C-source. Some Clostridia strains (*C. ljungdahlii*) can grow and form metabolites on syngas (a mixture of CO, CO_2 and H_2 produced during the gasification of coal or biomass). As an example, Evonik, using this procedure, developed a technical process for the production of pure 2-hydroxy-isobutyric acid, a starting material for the polymer poly-methacrylate (Plexiglass)

Biological feedstocks. Large quantities of biomass are produced on earth and might be used in the production of biofuel or chemicals. ~50 % of this biomass occurs in the oceans. Theoretically, this would be enough to cover both the food, energy and raw material needs of humanity. These resources are, however, quite unevenly distributed, and technologies for their sustainable use are still incomplete. In a sustainable global economy for the production of food, fuel and chemicals (bioeconomy), biotechnological processes and biorefineries (→330) are believed to play a key role.

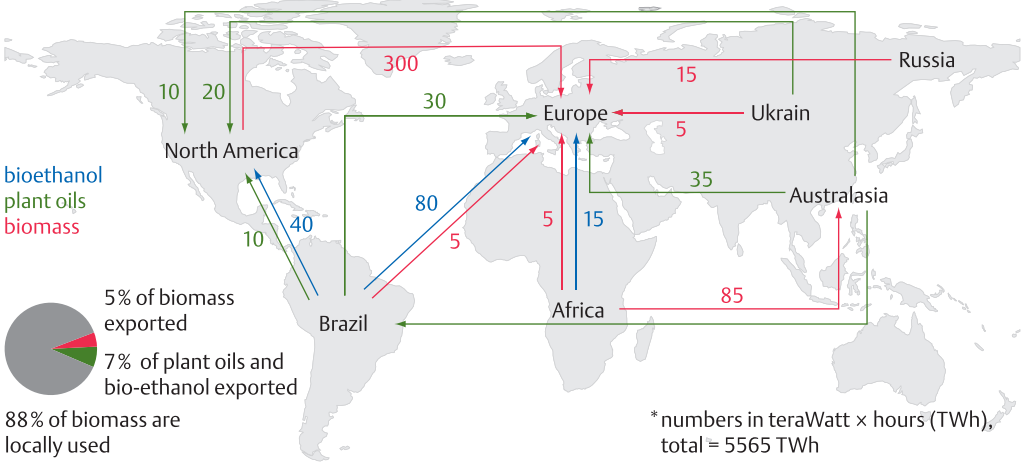
Direct use of biomass. (→82) Dry biomass such as straw from crops after harvest can be burnt in power stations and generate significant amounts of energy. Depending on the crop type, it may be costly to remove toxic or putrid gases.

Sugars from biomass. The most important processes investigated at present have the objective to transform biomass cellulose and hemicelluloses into fermentable sugars, in particular into glucose, xylose and arabinose. To this end, wood, harvest residues, etc., are crushed and decomposed in a first, usually physical step (e.. g., by high-pressure steam treatment) in order to release the polysaccharides from matrix components. In the next step, they are hydrolyzed by cellulase and hemicellulase enzymes.

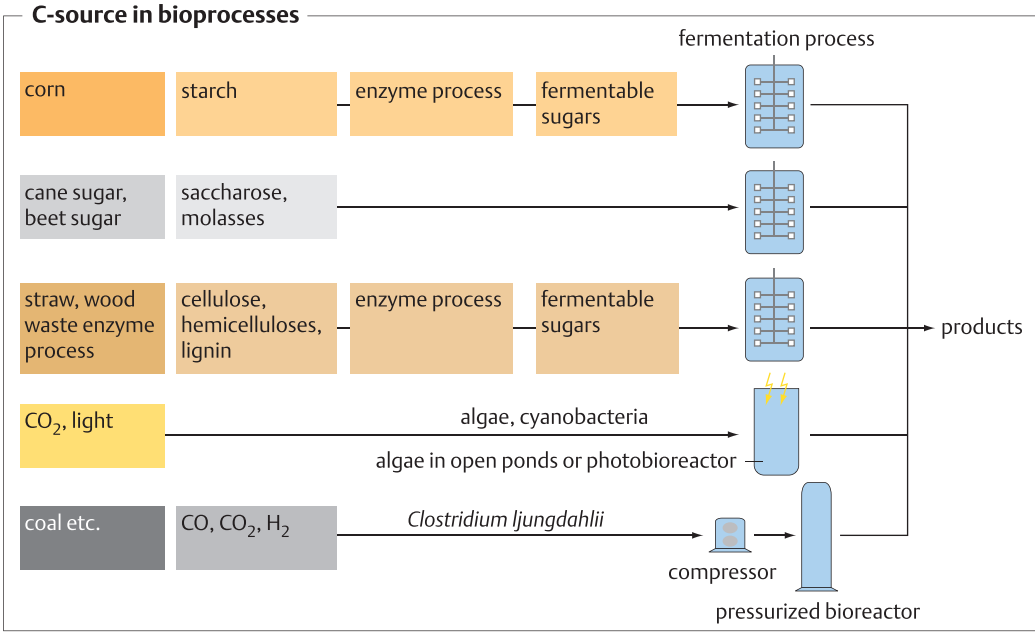
Lignin is an important builder of woody plants (20–30 % of dry mass, 20 billion t/y). Its biosynthesis is through radical polymerization of cinnamic acid (from phenylalanine) and thus contains a high proportion of aromatic compounds. Until now, attempts to use lignin for the sustainable production of aromatic chemicals or value-added products are limited.

Economic considerations. The petroleum-based economy in 2012 was based on 3.9 billion t of raw oil (C-content: 85 %). Transportation to the refineries was 60 % by >4000 oil tankers and 40 % in pipelines. 92 % of the raw oil was used for the generation of energy chemicals, and only 8 % was used for chemical synthesis. Biomass contains much less carbon and thus has less energy density. Its composition is diverse, and it is humid and biodegradable. As a consequence, and because of complex transport logistics (harvest cycles!), many local biorefineries are required to substitute oil refineries. In Germany, ~8000 biogas plants (→288) produce ~4000 MW electricity, which is the equivalent of just five coal-based power plants.

Biomass trade* (2020, estimated)



C-source in bioprocesses



Production of fossile raw materials and of biomass

raw material	annual production (Mill. t)	% content of carbon	main producer
crude oil	3,900	85.0	many regions
raw sugar	167	42.1	Brazil (42%)
palm oil	46	72.9	Malaysia (70%)
corn	700	44.0 (in starch)	USA (45%)
bagasse	530	50.0	Brazil (41%)
wheat straw	350	50.0	Russia (10%)
palm oil biomass	114	59.0	Malaysia (61%)
wood waste	900	50.0	Russia (17%)

Biorefineries

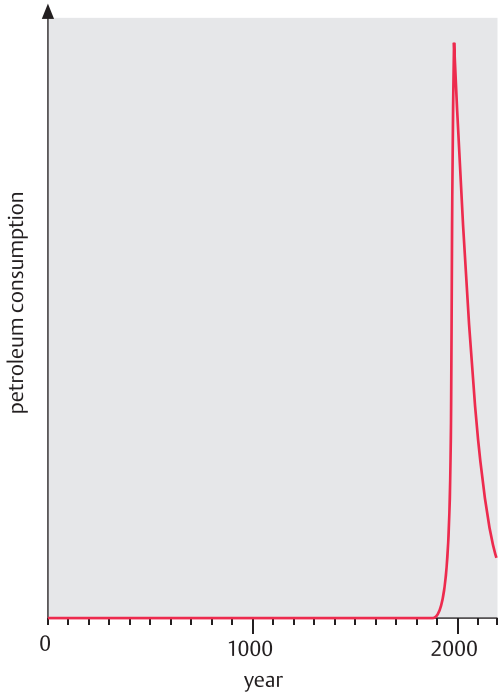
General. Technical chemistry has developed over the past 100 years into powerful integrated processes ("Verbund") for the economic production of energy chemicals and chemical intermediates from raw materials such as coal, gas or oil. The first step is usually done in a refinery (in the case of petrochemistry, by thermal cracking of longer-chain hydrocarbons) and, after fractionation, leads to energy chemicals such as gasoline, kerosine and diesel. In addition, building blocks for a wide range of chemicals are formed, which are grouped into C1- (from methane), C2- (from ethene), C3 (from propene), C4-, C5- and aromatic derivatives. In biorefineries, it is attempted to follow the same concept, but with sustainable raw materials such as CO_2 or biomass (from, e. g., wood, straw, waste materials, plant oils). Biotransformation (fermentation) plays a key role to generate valuable intermediates from these starting products. For this concept, the term "white biotechnology" is sometimes used.

Biorefineries. Typical raw materials for biorefineries are cellulose, hemicelluloses, lignins, starch and plant oils. Genetically modified plants which form less lignin, modified starch, or triglycerides of a desired chain-length may become important the future. Biorefineries are often distinguished according to their raw materials: 1) biorefineries using food crops such as corn, rye or wheat, 2) lignocellulose-based biorefineries using straw, reeds, wood, bagasse (from sugarcane production), or lignocellulose from papermaking, Lignocellulose-basierte Bioraffinerien (LCF), and 3) green waste biorefineries which use alfalfa, clover or grass as their raw material. As of 2014, about 200 biorefineries were under experimental operation. Preferred products were bioethanol, biobutanol and polymer building blocks such as lactic acid.

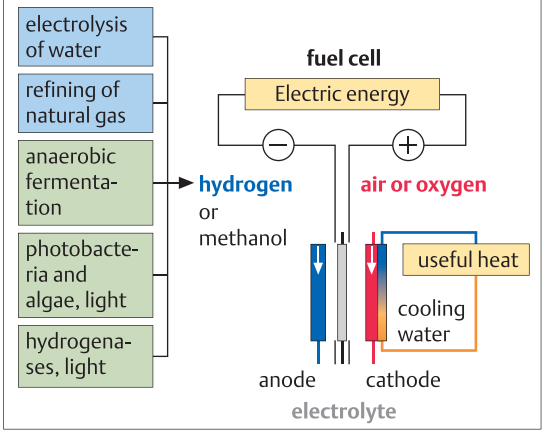
Bioenergy. Energy chemicals produced from sustainable raw materials are mainly bioethanol (→138), biodiesel (→162) and biogas (→288). Longer-range developments target biohydrogen for the operation of biological fuel cells. Most industrialized countries operate major programs for the production of bioethanol, to counter an expected shortage of oil. However, cheaper petroleum from fracking has, at least temporarily, curbed these programs. In the USA, three is a goal to increase by 2020 the share of bioethanol from the current 0.5 % to 10 % ("*roadmap to biomass technology in the US*"). Similar programs exist in the EU, in Japan and in China. *Biodiesel* is mostly produced by alkaline methanolysis of triglycerides such as plant oils or waste oils, e. g., "gutter oil," but can also produced from these materials via lipase catalysis. Biodiesel is already used as an additive to diesel fuel in cars. In the USA, China, France and Finland, manufacturing processes for special kinds of biodiesel are also certified "biokerosene" for aviation (2014). *Biogas* (a mixture of two parts of CH_4 with one part of CO_2) is formed by the anaerobic fermentation of biomass, and an established technology in sewage (anaerobic sludge digestion). It is also used in agriculture for a decentralized generation of energy. In Germany, >7,500 Biogas-plants were operational in 2012, and over 80,000 in China. In the long term, great expectations are placed on hydrogen as a fuel. Apart from its low weight, the major advantage of fuel cells is their higher energy conversion efficiency compared to traditional combustion engines. Thus, if biohydrogen could be economically produced through bio-photolysis of water (using alga or cyanobacteria), by photolytic degradation of biomass, by anaerobic bacterial fermentation from biomass, or even by a fuel cell using hydrogenase enzymes, such fuel cells would lead to a closed CO_2 carbon circle with enhanced energy efficiency. Such technologies are intensely studied throughout the industrialized world, but not yet mature enough for practical applications.

Ecoefficiency analysis. Chemical processes are analyzed today not only by economic, but also by ecological criteria. Parameters may vary in detail if used by different industries, but sustainability of raw materials plays as important a role as process yields, energy requirements and emissions (waste air, waste water, sludges). Toxic or inflammable auxiliary chemicals that must be used in a process create a handicap. If alternative manufacturing procedures for one product are compared using an ecoefficiency analysis, a clear preference for one out of several processes can usually be established. Biotechnological procedures often, but not always, are the most advantageous processes.

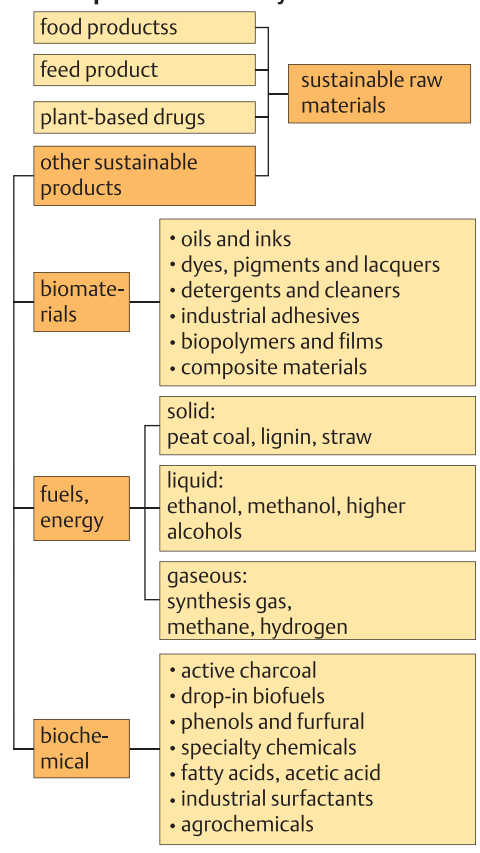
The petroleum age



Hydrogen technologies



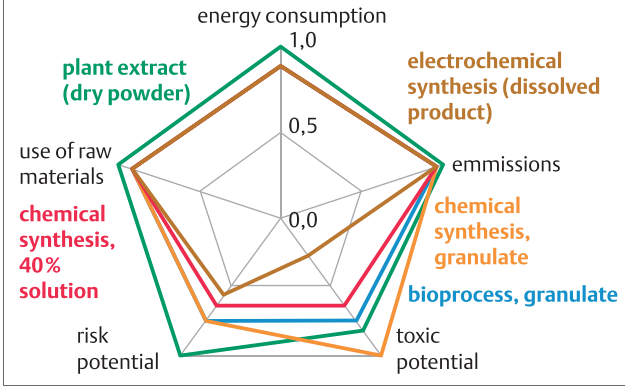
Concept of a biorefinery

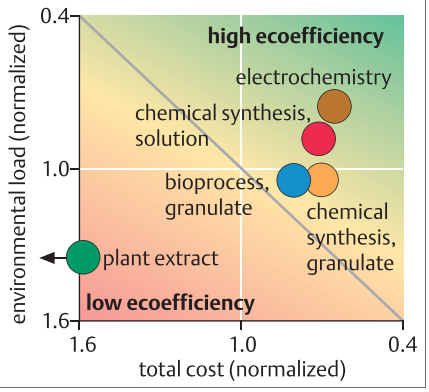


Cargill Plant in Blair, Nebraska

This biorefinery produces isoglucose, ethanol lactic acid, 3-hydroxypropionic acid and other chemicals from sustainable raw materials, e.g. corn and soja plants

Manufacture of indigo: ecoefficiency





Safety in genetic engineering

General. Early in the development of rDNA technology, scientists expressed concerns about the safety of the techniques employed to transfer genes from one organism to another. After a short moratorium, self-imposed by scientists in the USA and the UK after the Asilomar conference in 1975, safety in genetic engineering is now controlled in all industrialized countries by regulations, in some by laws. There are, however, significant differences in the details, e. g., in containment requirements.

United States NIH guidelines. Risk assessment is ultimately a subjective process. The investigator must make an initial risk assessment based on the risk group (RG) of an agent. Agents are classified into four risk groups according to their relative pathogenicity for healthy adult humans by the following criteria: 1) risk group 1 (RG1) agents are not associated with disease in healthy adult humans; 2) risk group 2 (RG2) agents are associated with human disease that is rarely serious and for which preventive or therapeutic interventions are *often* available; 3) risk group 3 (RG3) agents are associated with serious or lethal human disease for which preventive or therapeutic interventions *may be* available; 4) risk group 4 (RG4) agents are likely to cause serious or lethal human disease for which preventive or therapeutic interventions are *not usually* available.

European guidelines. In Europe, a statutory regulatory system is in force. It requires the assessment of the risks associated with the use of genetically modified organisms (GMOs). All work with living GMOs carried out within the European Union is regulated tightly by EU directives (European Directive on the contained use of genetically modified microorganisms (GMMs)). Each country implements the EU directives into its own legal system. In the UK the legislation covers all organisms, not only microorganisms. The Health and Safety Executive (HSE) operates and enforces legislation in Great Britain that controls the safety of activities involving genetically modified organisms in containment (Guidance on the genetically modified organisms [contained use] regulations 2000 and amendments). GMM are classified into 4 classes, depending on the risk assessment, with group 1 "unlikely to cause disease." When the assessed control measures fall between two containment levels, the activity must be classified at the higher level. The HSE must be notified of some activities involving modified organisms.

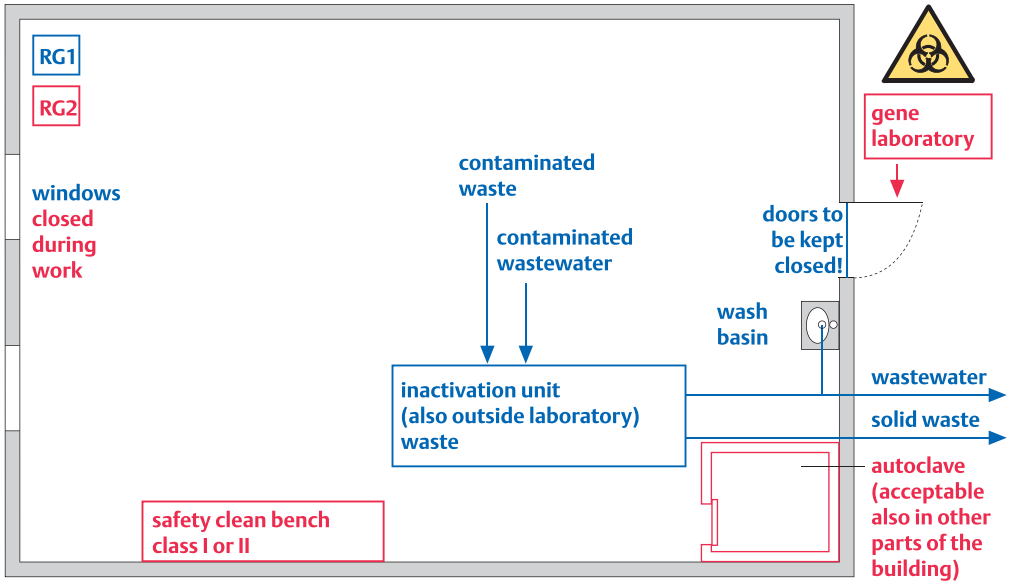
Cartagena protocol The *Cartagena Protocol on Biosafety to the Convention on Biological Diversity* is an international agreement which aims to ensure the safe handling, transport and use of living modified organisms (LMOs) resulting from modern biotechnology that may have adverse effects on biological diversity. It also takes risks to human health into account. It was adopted in 2000 and ratified in 2003.

Trained personnel. All staff who carry out genetic experimentation must have adequate professional training and must periodically receive additional formal instruction. A project leader (the principal investigator) is responsible for the professional conduct of all experiments and for their documentation. In some countries, a biological safety officer must be appointed for one or several laboratories who, in the name of the central executive officer of the institution, supervises the technical status of the laboratories and consults with the project leaders.

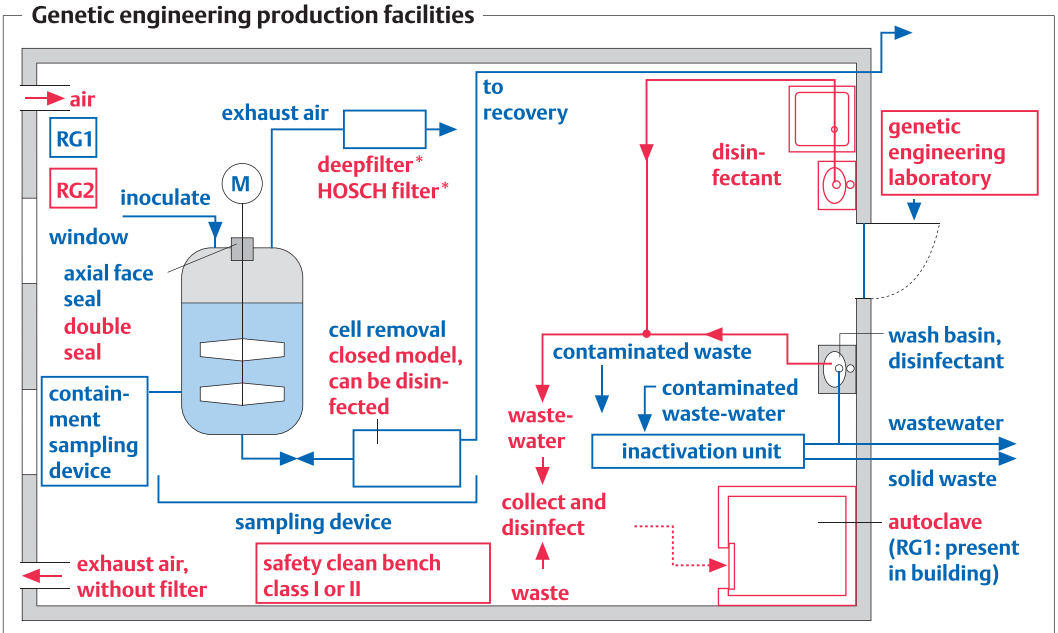
Laboratory equipment. Laboratories used for genetic engineering must first be cleared by the regulatory office in charge. Laboratories must be signposted, and admission is regulated. Depending on the experiments that are to be done, they must comply with different biosafety or containment levels (BSL), and construction features must meet these standards (e. g., clean benches with negative pressure, locks). The requirements for the discharge of wastewater, exhaust air, and waste increase with increasing safety levels. This also relates to health checks. In Germany, e. g., all employees working with RG2 agents or higher must undergo regular health checkups. Even stricter regulations are in place for recombinant product production facilities.

Documentation. Generally speaking, a regulatory office must be notified about the initiation of work at a genetic engineering laboratory, and the lab must be registered. All genetic experiments must be documented. In Germany, documents concerning RG1 agents must be kept for 10 years; those for RG2 agents and higher, for 30 years.

Genetic engineering laboratory RG1 and RG2 (according to German legislation)



Genetic engineering production facilities



*alternatively: cell separators, venturi washers, demister, foam control
*alternatively: sterilize exhaust air by heating

Risk assessment of genetically engineered microorganisms

	risk for man and environment	examples
RG1	no risk	lab strains of *E. coli*, bakers' yeast, transgenic plants and animals
RG2	low	some Pseudomonas strains, Xanthomonas
RG3	moderate	*Mycobacterium tuberculosis*, plant viruses
RG4	high	microorganisms strongly pathogenic to humans

Regulation of products derived from biotechnology

General. The manufacturing and sale of genetically engineered products is regulated by a complex set of rules intended to serve the safety of the consumer and the environment. In spite of differences in national regulations, these rules are similar in most countries. This survey is focused on US and EU regulations. It is appropriate to distinguish between regulations for pharmaceutical products obtained by recombinant technologies, and for food products or food additives.

Registration of a (bio)pharmaceutical. Pharmaceutical products are registered by the higher health authorities of a nation: in the USA this is the Food and Drug Administration (FDA), and in the European Union the European Medicines Agency (EMA) in London coordinates national examinations. In some developing countries, registration is done through the WHO. For registration of a drug, the following rules must be followed: 1) observance of the various GMP directives (Good Manufacturing Practice), in particular of the ICH directives, 2) comprehensive testing of
efficacy and safety, 3) precise documentation of the manufacturing process at certified sites (ISO 9001), and 4) documents concerning continuing quality control. The registration procedure usually starts with a preclinical phase, where efficacy and safety are investigated in the research laboratory and in animal experiments. If promising results are obtained, a process for manufacturing the drug is established and it receives the status of an experimental drug, entering a series of clinical experiments on humans. First, a small group of healthy volunteers is used to test drug safety (clinical phase I). This is followed by determination of efficacy and safety on groups of patients (clinical phase II). If all results are promising, testing of efficacy, safety, and side reactions is done with large groups of patients (clinical phase III). All data must then be submitted to the FDA or EMA, which can decide to register the drug or ask for further experiments. This process takes on average 11 years for chemical drugs, but just 9 years for bio-pharmaceuticals, at an average price exceeding 100 million US$ (only 1 among 5,000 preclinical candidates reaches registration). Parallel to the registration steps, the manufacturing process is standardized and documented. It must be proven that the product is free of chemical, microbiological, and genetic contaminants. This implies scrupulous control of the quality of air, water, starting materials, and storage conditions during manufacturing. In fermentation processes, the absence of biological contaminants (mycotoxins, retroviruses, etc.) must be ensured.

Food products affected by biotechnology methods or genetic engineering comprise transgenic fruits, vegetables, meat, fish or staple food such as flour, sugar and milk. In addition, food additives such as enzymes, dextrins, xanthan etc. may have been isolated from natural strains or produced by recombinant microorganisms. Since 1996, the US regulatory system does not require declaration in any of these cases. In the EU, on the other hand, novel food regulation requires the labeling of food products which have been modified by genetic engineering or which contain components prepared by gene technology.

Deliberate release of transgenic organisms. After > 20 years of tests, USA authorities usually permit the release of transgenic microorganisms, plants and animals to the environment after registration without further restraints. Manufacturing, sales and agricultural use of transgenic seed is permitted. In the European Union, transgenic plants have been approved since 2004, including food and feed products obtained from them. However, strict rules have been introduced both for release and for labeling of products obtained from genetically modified (GM) plants. Scientific evaluation of products is conducted by the European Food Safety Authority (EFSA) in Parma. The national rules usually follow the European regulations. Thus, the German Gene Law of 2005 requires the labeling and traceability of GM food and feed products, and limits the use of antibiotic resistance genes to 10 years.

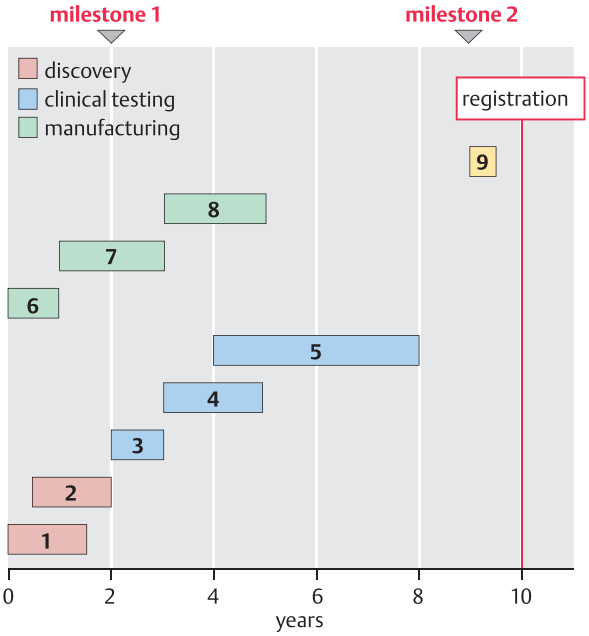
The Cartagena Protocol *on Biosafety* is an international agreement which aims to ensure the safe handling, transport and use of living modified organisms (LMOs) resulting from modern biotechnology that may have adverse effects on biological diversity or pose risks to human health. It was adopted in 2000 and ratified in 2003.

Germany: pharma pipeline 2013 (clinical tests phase I–III)

antibodies and antibody fragments	333	due to faster registration processes, there is a general trend to biopharmaceuticals
recombinant proteins	122	
vaccines	111	
gene therapy agents	21	

in 2013, 215 biologics were registered in Germany

Registration of a pharmaceutical



1 discovery, efficacy
2 animal tests
3 clinical testing phase I: safety for healthy humans (20–30 volunteers)
4 clinical testing phase II: efficacy and safety (100–300 patients)
5 clinical testing phase III: large-scale experiments in hospitals on efficacy, safety and side effects (1000–5000 patients)
6 engineering of plant
7 construction
8 manufacturing procedure under GMP
9 inspection of plant
MS1 registration as experimental drug
MS2 submission of all documents on testing and manufacturing to registration office

patent protection ceases after 20 years

Criteria for registration

efficacy	effects are proven and of national economic relevance*
safety	based on animal and clinical experiments, side effects can be estimated and are low compared to efficacy
manufacturing safety	the manufacturing process is standardized and follows GMP rules. The preparation is free of side products, pyrogens, viruses, bacteria, and infectious DNA
other	price relates to benefit

*except for orphan drugs (drugs active against rare diseases)

Enzymes as feed or food additives

Europe	Feed enzymes must be approved by the European Food Safety Authority (EFSA) in Parma. On the basis of dossiers containing detailed data on manufacturing, health and safety aspects as well as performance in animal feed. A similar process for approval has een established for food enzymes (FIP) where a differentiation is made between classical and GMO-derived enzymes. Regulation depends on whether the enzyme is considered an additive or a processing aid.
USA	Food ingredients, including enzymes, may either be FDA-approved food additives or GRAS (generally recognized as safe) ingredients exempt from food additive status through a notification to the FDA (Food and Drug Administration), http://www.fda.gov . A food enzyme with GRAS status must be recognized as safe by qualified experts through scientific procedures for its intended use.

Ethical considerations and acceptance

General. Various aspects of genetic engineering and cell biology have raised science-philosophical and ethical questions. These include 1) how to protect a person's genetic information, 2) how to justify gene therapy, in particular of haploid (germline) cells, leading to inheritable traits, 3) the cloning of human embryos, 4) stem-cell therapy, and 5) animal welfare, e. g., knockout animals used in drug development and biopharmaceutical production in transgenic animals, and 6) the military use of biotechnology and genetic engineering. Although modern biotechnology is socially accepted in principle, it is widely held that benefits must be balanced against risks.

Individual genetic information. The advantages of genetic screening, in particular after having completed the sequencing of thousands of human genomes, make it reasonable to believe that individual dispositions toward genetic diseases will become predictable in the future, e. g., by prenatal diagnostics (→302). This raises the question of how doctors and society should handle this information (e. g., should preimplantation genetic diagnosis (PGD) be allowed, and should abortion be accepted if an embryo displays a genetic disposition, but not a deterministic risk, of developing an incurable disease?). We also need to clarify to what extent governmental bodies may store individual genetic data for, e. g., the purpose of faster screening in criminal investigations, or if insurance companies and employers may access these data to balance insurance or employment vs. health risks.

Gene therapy. (→304) In a democratic society, somatic gene therapy depends on the informed consent of the patient, and risks are low. However, if gene therapy is applied to haploid eggs or sperm, modified genetic patterns may be inherited by the offspring without their consent. The necessary techniques are rapidly developing in animal experiments, but their acceptance and success in human therapy is far from clear.

Genetic manipulation. The use of microorganisms in the production or transformation of fine chemicals is hardly debated, even if genetic manipulations are involved. Concerns about genetically engineered plants (→282) are usually confined to questions of whether the food products obtained from such plants are safe, and if the long-term ecological consequences of their cultivation can be controlled. In comparison, genetic manipulation of animals raises more public concern. The goals of transgenic animal experiments (production of biopharmaceuticals, use of knockout animals in drug research) (→270) play a minor role in these concerns.

Animal and human clones. Following the production of identical clones in sheep, mice, pigs, cattle, and goats, the reproductive cloning of humans has met with major ethical discussions. These include the population genetic consequences of techniques such as the sexing of progeny or in-vitro fertilization and embryo transfer.

Stem cell therapy. A main discussion point is at which stage of development human life has started and must be protected and if therapeutic alternatives can be developed. The advent of iPS technology may solve such controversies.

Military or terrorist use. Standard biotechnological procedures can be used to produce agents of warfare, such as *Bacillus anthracis* spores or smallpox virus. Genetic engineering techniques could be used to develop bioweapons that overrun the defensive mechanisms of the immune system. Although bioweapons have been outlawed by the Geneva Convention, several nations are suspected of continuing production of "B weapons".

Public acceptance. In the USA, which is at the forefront of most developments in genetic research and gene technology (⅔ of all gene therapy experiments are done in the USA), the public generally has fewer reservations compared to Europe and Japan. In most nations, the medical use of biotechnology and genetic engineering meets with public support. The production of transgenic animals meets with considerable public criticism in Europe, al-though in practice transgenic mice have become a standard tool for biomedical research. Transgenic plants meet little enthusiasm in European nations, but are standard crops in the USA and Canada and of increasing importance in nations like China. Genetically modified enzymes are widely used in technology and food technology, without major public opposition.

Science-philosophy arguments concerning genetic engineering

categorical argument

Some human activities such as genetic engineering are fundamentally reprehensible. Developing this technology, "man plays God" and claims competencies beyond his capacities, degrading nature to the course of his technical manipulations.

pragmatic argument

The key objective of genetic engineering is to reduce the suffering of diseased individuals. The procedures which are applied must, however, be safe, and the patient must be able to decide if he or she wishes to apply genetic diagnosis or therapy.

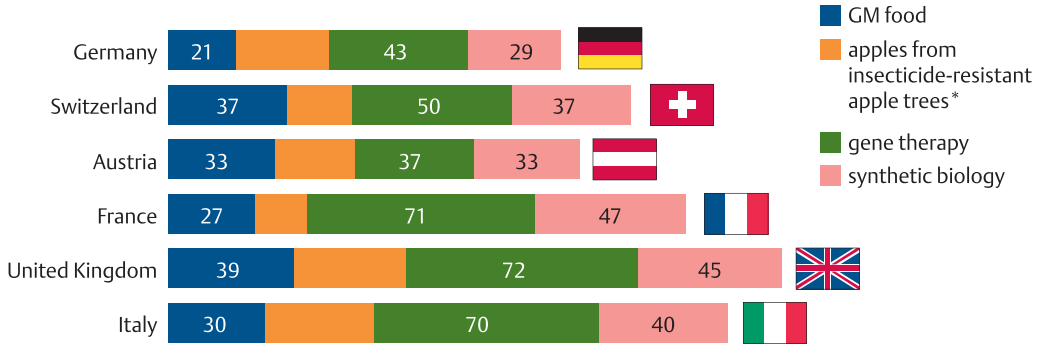
social policy argument

The social effects of genetic engineering cannot be estimated. In genetic therapy, wrong priorities are chosen, better prophylaxis would be more desirable. We start down a slippery slope that will lead us involuntarily to inhumane practices towards the next generations ("eugenics bottom up")

Problematic areas of genetic research

topic	state of the art	regulation or trend
cloning of humans	cloning of animals possible	not permitted
use of embryonic stem cells	growing expertise	permitted, but regulated
artificial insemination, sexing, surrogate mothers	state of the art in animals	artificial insemination permitted, sexing and surrogate mothers forbidden
prenatal diagnosis	cytological methods established, DNA-based diagnosis partially established	permitted, abortion permitted after medical indication
identifying genetic risks by genetic screening	possible for some monogenic diseases	under debate if one gene defect is predictive and if diagnosis is acceptable for incurable diseases; strict data protection required towards employers, insurance companies
knockout animals for drug research	widely established	generally accepted, but hotly debated by animal protection groups
food and biopharmaceutical production using transgenic animals or plants	many techniques established	debated in view of consumer protection, animal protection, ecological consequences
transgenic microorganisms or cell lines for production of biopharmaceuticals	established	widely accepted

Acceptance in European nations (2010)



*the answers were more positive if genes from insect-resistant apple trees were transferred

Patents in biotechnology

General. The governments of all industrialized countries can, through their patent offices, grant inventors a monopoly for a limited period of time. Generally such a monopoly is restricted to a period of 20 years. There are different kinds of intellectual property protection, such as patents or utility models on the commercial use of the manufacture of materials, processes for their manufacture, and the uses thereof. For a patent to be granted the invention must be novel, inventive, and applicable. Therefore, in *patents on products or compounds* the material described in the invention, such as a chemical compound, mixture, or a piece of equipment must be novel, inventive, and of economic value. The same applies to *process patents.* In addition to these three prerequisites the invention has to be disclosed in such a manner that a person skilled in the art can practice it without undue burden. Discoveries, scientific theories, mathematical methods, and aesthetic creations are not patentable. Neither are software or doing business as such, nor therapeutic, diagnostic, or surgical methods for humans and animals patentable. However, pharmaceutical compounds or mixtures are patentable. Intellectual property rights may not be granted on inventions that would be contrary to "ordre publique" or morality.

Patenting process. Any natural or legal person can file a patent application with the national or regional patent office. It is advisable, however, to enlist the help of a patent attorney, both for writing the application and for dealing with the patent office. In view of competing inventions, is important to obtain a priority date as early as possible. The priority date in most countries is the day of submission of the application to the patent office. Within 12 months after submission, an application based on the first submission can be filed as an international (PCT) or regional patent application (e. g., to the EU patent office). The text of the application is published by the patent office 18 months after priority deposition. Examination of the patent by the patent office always depends on a request for examination by the applicant, often 4 to 5 years after deposition. A patent is valid only after patent issuance, following examination. After the patent has been granted, third parties can oppose the patent immediately, as well as during the lifetime of the patent by an action of annulment if the patent is considered not novel, not inventive or not applicable. The priority date in most countries is the date of submission of the application to the patent office ("first to file"), but in the USA it is the date of documentation of the invention ("first to invent"), for example, evidence of the invention in a laboratory journal. In the USA and Japan, patents can be submitted up to 12 and 6 months, respectively, after an invention has been published ("period of grace"); under the European Patent Convention, any kind of prior publication prevents patenting. Patents, once granted, remain valid for 20 years after the filing date if the maintenance fees are paid but can be prolonged for 5 years in some countries for, e. g., new pharmaceuticals. The annual costs for a patent application range between € 1,000 and 20,000 or more, and its prosecution, maintenance and enforcement are very expensive. To prevent high translation costs for registration with foreign patent offices, a patentee can file an application according to the PCT treaty in one of seven official languages. After he has requested international examination, the patentee can decide within up to 30 months after the priority date whether to pursue the application in any of the more than 110 PCT contract states.

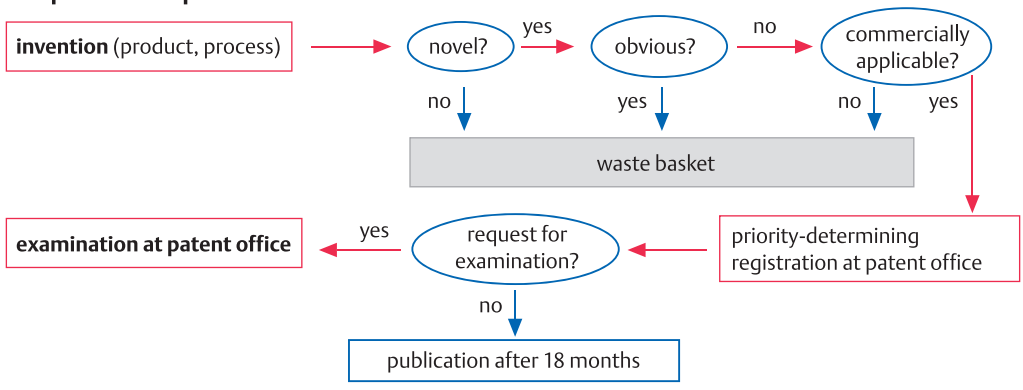
Patents in biotechnology. Materials isolated from living beings or manufactured by biotechnological processes are patentable under the rules mentioned above. Living beings especially bred for economic purposes are also patentable, e. g., microbial production strains or cell lines, transgenic plants and animals, but not germplasms. The requirement for "enablement" necessary for a person skilled in the art to carry out the invention may often require deposition of a sample of the organism necessary for the manufacture of a product with a depositary institution (Budapest treaty, 1965). For claims concerning proteins or DNA-based products, however, indication of the amino acid or nucleic acid sequence is usually enough. Mutagenized genes or proteins or polymorphic DNA may be patented. Gene sequences of expressed sequence tags (ESTs) are not patentable without revelation of their specific properties. Transgenic plants and animals are patentable depending on the specific details.

Patent categories and their relation to biotechnology

examples from biotechnology

substance patents		process patents	
substances	cloned genes, recombinant proteins, monoclonal antibodies, plasmids, promoters, vectors, cDNA sequences, monovalent vaccines	processes	DNA isolation, DNA synthesis, preparation of vectors, purification processes for proteins
substance mixtures	polyvalent vaccines, bioinsecticides, pharmaceutical preparations, microorganisms, transgenic plants and animals	process protocols	hybridization assays, diagnostic procedures, PCR methods, analysis of mutants
equipment	pulse-field gel electrophoresis, DNA sequencer, biolistics gene cannon	applications	applications of bioinsecticides, fermentation protocols for genetically modified microorganisms

Steps toward a patent

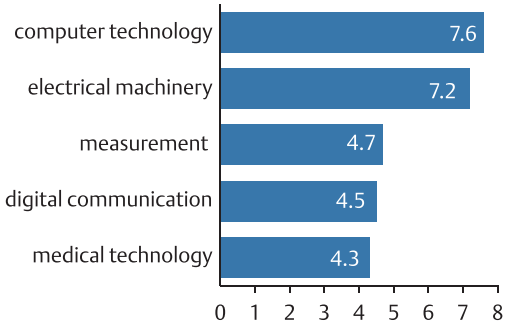


patent offices		
USA	US Patent Office, Washington	www.uspto.gov/
Europe	European Patent office, Munich	www.european-patent-office.org/cn.php
Japan	Japan Patent Office, Tokio	www.jpo.go.jp/
PR China	State Intellectual Property Office	www.sipo.gov.cn
WIPO	World Intellectual Property Organization, Genf	www.wipo.int/portal/index.html.en

Patent statistics (WIPO, 2014)

patents in force 2013, world	9.45 Mill.
of which US patents	2.39 Mill.
of which Japanese patents	1.79 Mill.
patent applications 2013, world	~1.17 Mill.
issued patents 2013, world	2.57 Mill.
of which China	~825,000
(only applications abroad)	~30,000
of which USA	~572,000
of which Japan	~328,000
of which European Patent Office	~149,000

most published applications 2013 (%):



International aspects of biotechnology

General. Like most technical and scientific breakthroughs in human history, biotechnology and genetic engineering have resulted from a mosaic of many national contributions. Often, quantum jumps in basic research or an economic crisis have helped to move a step ahead; this is certainly true for the development of fermentation industries in Europe and Japan. Genetic engineering, however, is clearly a brainchild of the USA. Since ca. 1971, a growing number of academic entrepreneurs in this country have founded companies, which in many cases attracted private venture capital, to be finally sold to larger pharmaceutical companies or to grow into an independent larger company. In the meantime, there are > 2500 venture-capital based biocompanies in the USA and in Europe, focused on biotechnology, genetic engineering, and bioinformatics.

USA. The first companies to be founded around molecular genetics were Cetus (1971), Genex (1972), Genentech (1977), Biogen (1978), and Amgen (1980). The initial incentive was usually a new technology that had been secured by patents (e. g., for Genentech, the heterologous expression of genes according to the Cohen-Boyer patent). Some of these companies succeeded, while still funded by venture capital, in inventing and patenting a large number of innovations within just a few years. Genentech obtained patents on the production of human insulin in *E. coli* (1976) and of human factor VIII (1978) and human TPA (1980) in animal cells. License fees from these patents and, to some extent, their own marketing efforts led to a significant increase in turnover and stock value. In 1990, Roche, Switzerland, acquired 60 % of the Genentech shares for 2.1 billion US$; shortly after, it also acquired Cetus with its global patent on PCR. However, as of today, the number of successfully operating startup companies is still small, compared to the invested capital. Since ~2000, it is predominantly the public capital markets which are drivers for the development of new technologies such as immunotherapy, gene therapy and stem cell research. Large pharmaceutical companies such as Novartis, Roche, Aventis, GlaxoSmithKline, Merck, Bayer, Pfizer, Bristol-Myers Squibb, etc., all operate transnational research and development centers, to use the large potential of genomic and post-genomic research for the development of novel drugs fitting into their marketing strategies. In a continuous series of mergers and acquisitions, they take over smaller competitors or start-up companies, adjusting their portfolios or buying into new drugs or technologies at early phases.

Europe. In the beginning of genetic engineering, most companies in continental Europe were slow to set up their own research laboratories in this field, and venture companies also developed slowly. Even in 2014, Europe did not catch up with the USA, though a growing number of successful small and medium enterprises have been established in most European countries, with the UK and Germany sharing the top position. The European Union is strongly supporting biotechnology-related programs both by funding and by enhancing the mobility of young researchers.

Japan. Venture companies are rare, but nearly all large companies have strong stakes in genetic engineering and cell biology. After the privatization of the universities and public research institutes such as RIKEN, the number of venture companies increased and was around 500 in 2011. Only a few companies that have diversified into biotech from other business have been successful. Examples are Ajinomoto, Takara Bio and Kirin Beer.

China has taken great efforts to develop modern biotechnology. Institutes of the Chinese Academy of Sciences (CAS) and universities such as Peking, Tsinghua or Jiaotong University are catching up with internationally leading groups, and the number of patents and publications is now second only to the USA.

Innovation and performance. Most of the innovations that help to advance biotechnology originate in the industrialized nations of Europe, Japan, increasingly China, but above all, the USA. If scientific publications in key areas such as molecular genetics and cell biology are taken as an indicator, the US contributes about 40 % of all global communications.

Governmental programs. As one of the megatechnologies of the 21st century, the development of biotechnology is supported in all industrialized nations, and also in many developing nations such as China. The development of core competencies and the training of personnel both play important roles. A key task is to train innovative young minds for the challenges of new multidisciplinary technologies that involve biology, chemistry, ecology, engineering, and informatics, and last but not least, economics.

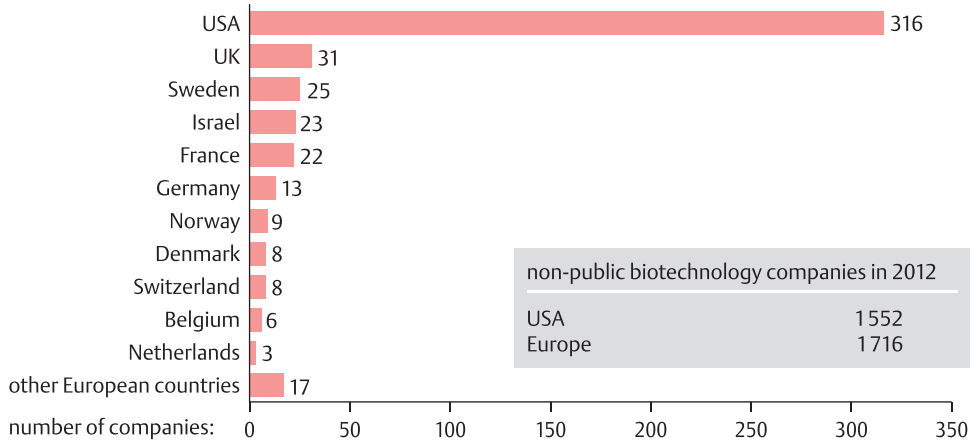
International comparison of performance - citations in PubMed, 2013

	world	USA	Germany	UK	France	Japan	China
molecular genetics	62,251	24,765	4,219	3,771	2,647	6,480	2,529
fermentation	9,526	870	653	155	141	1,346	1,300
pharmaceuticals	37,361	13,558	3,099	2,295	1,730	3,882	1,300

Performance of biotech companies

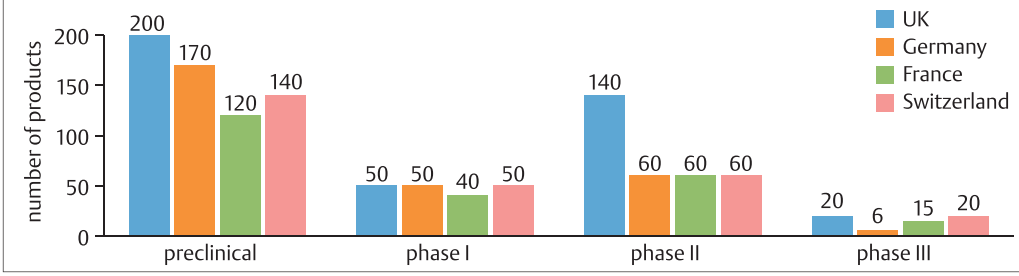
	USA [US$]			Europe [€]		
	2000	2003	2012	2000	2003	2012
total number of companies	1,379	1,473	1,859	1,570	1,861	1,799
of which are public companies	339	314	316	105	96	165
employees (× 1,000)	174	198	~ 100	61	78	~ 52
R & D expenditure (in billion)	14.7	17.9	19.3	5.0	6.3	4.9
sales (in billion)	26.6	39.2	63.7	8.7	11.2	20.4
market capitalization (in billion)	6.2	5.4	360	1.6	1.9	~ 80

Public biotechnology companies, USA and Europe (2012)



The market capitalization of the US-companies in 2012 was about 360 billion US$, that of all European biotech companies about 80 billion US$

Biotech development products (2012)



Financing of biotechnology companies (2012, in million US$)

	USA			Europe		
	2000	2003	2012	2000	2003	2012
IPOs	6,698	448	765	2,950	–	40
follow-on capital	23,237	11,131	6,620	2,447	1,602	948
venture capital	3,207	2,826	4,126	1,154	1,040	1,934
total	33,142	14,405	63,279	6,551	2,642	4,164

Further Reading

For this book, a large number of informations from books, book sections, reviews, publications and personal communications were used. It is impossible to cite all these sources.
In view of the presumed interests of the reader, I have chosen to provide a limited selection of about 600 citations, which are ordered according to the entries of this book. They include also a limited number of highly relevant citations from earlier editions.
In addition, the following review journals were preferentially used:

Kirk-Othmer Encyclopedia of Chemical Technology, 4th Edition, Wiley Online Library
Ullmann's Encyclopedia of Industrial Chemistry, 5th Edition, Wiley-VCH
Advances in Biochemical Engineering/Biotechnology, Springer
Advances in Biotechnological Processes, Wiley Online Library
Applied Microbiology and Biotechnology, Springer
Biochemical Society Transactions, Portland Press
Biotechnology – A Comprehensive Handbook, 2nd Edition, Wiley-VCH
Biotechnology and Bioengineering, Wiley Online Library
Chemical & Engineering News, American Chemical Society
Current Biology, CellPress
Current Opinion in Biotechnology, Current Biology Publications
Current Opinion in Immunology, Current Biology Publications
Current Opinion in Microbiology, Current Biology Publications
Current Opinion in Structural Biology, Current Biology Publications
Current Opinion in Genetics and Cell Biology, Current Biology Publications
Journal of Biotechnology, Elsevier Publishers
Nature Biotechnology, Nature Publishing Group
Trends in Biochemical Sciences, Elsevier Publishers
Trends in Biotechnology, Elsevier Publishers
Trends in Cell Biology, Elsevier Publishers
Trends in Microbiology, Elsevier Publishers

The internet has become a most valuable resource for scientific information. Wikipedia, in particular, is a treasure trove for all kinds of reviews as well as for highly specialized information. Wikipedia articles are cited as, e. g., Wikipedia: "Interferon". Since there is no guarantee for continuous access to a website, however, other citations of internet articles were done in a conservative manner, and movies, e. g., from YouTube, some of them quite excellent, were omitted.

General Textbooks and Compendia

Atkinson, B., Maviura F. (1991) *Biochemical Engineering and Biotechnology Handbook*, 2nd Edition, Mamillan Publishers, ISBN 1-56159-012-6
Demain, A., Davies J. (1999) *Industrial Microbiology and Biotechnology*, 2nd ed., ASM Press, ISBN 1-55581-128-0
Walsh, G. (2002) *Protein Biotechnology and Biotechnology*, John Wiley & Sons, ISBN 978-0471899075
Rehm, H. J., Reed, G., Pühler, A., Stadler, P. eds., *Biotechnology – A Multi-Volume Comprehensive Treatise*, 2nd Edition. Wiley-VCH, ISBN 1-56081-602-3
Brown, T. A. (2010) *Gene Cloning and DNA Analysis: An Introduction*, 6th edition, Wiley e-text, ISBN 9781405181730
Soetaert, W., Vandamme, E. J., Demain, A. L. eds., (2010) *Industrial Biotechnology: Sustainable Growth and Economic Success*, Wiley VCH, ISBN 978-3-527-31442-3
Thiemann, W. J., Palladino, M. A., Cummings, B. (2014) *Introduction to biotechnology*, 3rd ed. Pearson Education, ISBN 9781292027616
Sahm, H., Antranikian, G., Stahmann, K. P., Takors, R. eds. (2013) *Industrielle Mikrobiologie*, Springer Spektrum, ISBN 978-3-8274-3039-7
Renneberg, H., Berkling, V. (2013) *Biotechnologie für Einsteiger*, 4. Aufl., Springer Spektrum, ISBN 978-8274-3047-2
Renneberg, R. (2008) *Biotechnology for Beginners*, Academic Press, ISBN 978-0-12-373581-2

Microbiology 6–25

Fuchs, G., Schlegel, H. G. (2006) *Allgemeine Mikrobiologie*, Georg Thieme Verlag, 978-3-13-444608-1

Madigan, M. T., Martinko, J. M. (2006) *Brock Biology of Microorganisms*, 11. ed., Pearson Prentice Hall, ISBN 0-13-196893-9

Black, J. G., Black, L. J. (2015) *Microbiology: Principles and Explorations*, 9th ed., Wiley, ISBN 978-1-118-74316-4

Viruses 6/7

http://en.wikipedia.org/wiki/Virus

Zinkernagel, R. M. (1996), Immunology taught by viruses. *Science* 271, 173. PMID: 8539616

Miyazaki, Y., Fujita, M., Nomaguchi, M., & Adachi, A. (2012). Structural biology for virus research. *Frontiers in Microbiology*, *3*, 91. doi:10.3389/fmicb.2012.00091

http://www.ictvonline.org/index.asp

http://www.ncbi.nlm.nih.gov/genome/viruses/

Bakteriophages 8/9

http://en.wikipedia.org/wiki/Bacteriophage

Ackermann, H. W. (2012) Who went into phage research? Bacteriophage. 2012 Jan 1;2(1):55–59. PMID: 22666657

http://www.ebi.ac.uk/genomes/phage.html

Microorganisms 10/11

http://en.wikipedia.org/wiki/Microorganisms

http://www.microbeworld.org

Black, J. G., Black, L. J. (2015) *Microbiology: Principles and Explorations*, 9th ed., Wiley, ISBN 978-1-118-74316-4

Bacteria 12/13

Lengeler, J., Drews, G., Schlegel, H. (1999) *Biology of the Prokaryotes* Thieme, ISBN 3-13-108411-1

http://en.wikipedia.org/wiki/Bacteria

http://www.ncbi.nlm.nih.gov/genomes/MICROBES/microbial_taxtree.html

Yeasts 14/15

http://en.wikipedia.org/wiki/Yeasts

http://www.yeastgenome.org

http://www.pombase.org

Fungi 16/17

http://en.wikipedia.org/wiki/Fungi

Johnson, E. A. (2013). Biotechnology of non-Saccharomyces yeasts – the ascomycetes. *Applied Microbiology and Biotechnology*, *97*(2), 503–17.

Johnson, E. A. (2013). Biotechnology of non-Saccharomyces yeasts-the basidiomycetes. *Applied Microbiology and Biotechnology*, *97*(17), 7563–77.

http://www.cbs.knaw.nl/Collections/

http://www.indexfungorum.org

Algae 18/19

http://en.wikipedia.org/wiki/Algae

http://www.algaebase.org

Jones, C. S., & Mayfield, S. P. (2012). Algae biofuels: versatility for the future of bioenergy. *Current Opinion in Biotechnology*, *23*(3), 346–51.

Hariskos, I., & Posten, C. (2014). Biorefinery of microalgae – opportunities and constraints for different production scenarios. *Biotechnology Journal*, *9*(6), 739–52.

http://www.algaebase.org

http://www.shigen.nig.ac.jp/algae/

Some bacteria of importance for biotechnology 20/21

http://en.wikipedia.org/wiki/Escherichia_coli

http://en.wikipedia.org/wiki/Bacillus

http://en.wikipedia.org/wiki/Pseudomonas

http://en.wikipedia.org/wiki/Corynebacterium

http://en.wikipedia.org/wiki/Streptomyces

http://www.ncbi.nlm.nih.gov/genome/browse/

Dixon, B. (1996) *Power unseen: How Microbes rule the World*, W. H. Freeman, ISBN 0-7167-4550-X

Blattner, F., Plunkett, G., Bloch, C., Perna, N., Burland, V., Riley, M., Collado-Vides, J., Glasner, J., Rode, C., Mayhew, G., Gregor, J., Davis, N., Kirkpatrick, H., Goeden, M., Rose, D., Mau, B., Shao, Y. (1997), The Complete Genome Sequence of Escherichia coli K-12 *Science* 277, 1453

Sahm, H., Antranikian, G., Stahmann, K. P., Takors, R. eds. (2013) *Industrielle Mikrobiologie,* Springer Spektrum, ISBN 978-3-8274-3039-7

Zahoor, A., Lindner, S. N., Wendisch, V. F. (2012). Metabolic engineering of Corynebacterium glutamicum aimed at alternative carbon sources and new products. *Computational and Structural Biotechnology Journal*, 3, e201210004

Wieschalka, S., Blombach, B., Bott, M., Eikmanns, B.J. (2013). Bio-based production of organic acids with Corynebacterium glutamicum. *Microbial Biotechnology*, 6(2), 87–102.

Vertès, A.A., Inui, M., Yukawa, H. (2012). Postgenomic approaches to using corynebacteria as biocatalysts. *Annual Review of Microbiology*, 66, 521–50.

Juhas, M., Reuß, D., Zhu, B., Commichau, F.M. (2014). Bacillus subtilis and Escherichia coli essential genes and minimal cell factories after one decade of genome engineering. *Microbiology (Reading, England)*, 160(Pt 11), 2341–51.

Microorganisms: isolation, preservation, safety 22/23

http://en.wikipedia.org/wiki/Biosafety_level

Microorganisms: strain improvement 24/25

Crueger, A. (1993) *Mutagenesis*, in Biotechnology Second, Completely Revised Edition, 5. Vol. 2, Pühler, A. (ed.). VCH Verlagsgesellschaft, ISBN 3-527-28312-9

Biochemistry 26–37

Lehninger, A.L., Nelson, D.L., Cox, M.M. (2008), *Principles of Biochemistry*, 5th Edition, Worth Publishing, ISBN 978-1429208925

Stryer, L., Berg, J.M., Tymoczko, J.L. (2002), *Biochemistry*, 4th Edition, W. H. Freeman and Co, ISBN 978-0123735812

Bmunk, K. (Hrsg) (2008), Taschenlehrbuch Biologie Biochemie Zellbiologie, Georg Thieme Verlag, ISBN 978-3-13-144831-6

Amino acids, peptides, proteins 28/29

http://en.wikipedia.org/wiki/Amino_acid
http://en.wikipedia.org/wiki/Peptide
http://en.wikipedia.org/wiki/Protein
http://en.wikipedia.org/wiki/Protein_structure
http://www.rcsb.org/pdb/home/home.do

Enzymes: structure, function, kinetics 30/31

http://en.wikipedia.org/wiki/Enzyme
http://en.wikipedia.org/wiki/Enzyme_kinetics
http://www.brenda-enzymes.org
http://www.chem.qmul.ac.uk/iubmb/enzyme/

Sugars, glycosides, polysaccharides 32/33

http://en.wikipedia.org/wiki/Carbohydrate
http://en.wikipedia.org/wiki/Glycoside
http://en.wikipedia.org/wiki/Polysaccharide

Lipids, membranes, membrane proteins 34/35

http://en.wikipedia.org/wiki/Lipidomics
http://en.wikipedia.org/wiki/Biological_membrane
http://en.wikipedia.org/wiki/Cell_membrane
http://en.wikipedia.org/wiki/Membrane_protein
http://en.wikipedia.org/wiki/Lipid_bilayer

Metabolism 36/37

Michal, G., Schomburg, D. (2012) *Biochemical Pathways* 2nd ed., J. Wiley & Sons, ISBN 978-0-470-14684-2

Genetic Engineering 38–75

Watson, J.D. (1965) *Molecular Biology of the Gene*, W.A. Benjamin, Inc., Library of Congress 65-19424

Watson, J., Gilman, M., Witkowski, J., Zoller, M. (1992) *Recombinant DNA*, 2nd ed., Scientific American Books, ISBN 0-7167-2282-8

Nicholl, S.T. (2008) *An Introduction to Genetic Engineering*, 3d ed., Cambridge University Press, ISBN 978-0521615211

Brown, T.A. (2010) *Gene Cloning and DNA analysis*, Wiley-Blackwell, ISBN 978-1405181730

Dale, J.W., von Schantz, M., Plant, N. (2012) *From genes to genomes*, Wiley Blackwell 3rd ed., ISBN 978-0-470-68386-6

DNA: structure and function 38–41

http://en.wikipedia.org/wiki/DNA_structure

Rehmann, J. (1996) *Nucleic Acids*, in Kirk-Othmer Encyclopedia of Chemical Technology 4th Edition 17, 507. Wiley Interscience, ISBN 0-471-52686-X

Götz, F. (1993) *Structure and Function of DNA* in Biotechnology Second, Completely Revised Edition, 191. Vol. 2, Pühler, A. (ed.). VCH Verlagsgesellschaft, ISBN 3-527-28312-9

RNA 42/43

http://en.wikipedia.org/wiki/RNA

http://en.wikipedia.org/wiki/Systematic_evolution_of_ligands_by_exponential_en richment

http://en.wikipedia.org/wiki/Cell-free_protein_synthesis

http://en.wikipedia.org/wiki/Small_interfering_RNA

http://en.wikipedia.org/wiki/Aptamer

Rosenblum, G., Cooperman, B. S. (2014). Engine out of the chassis: cell-free protein synthesis and its uses. *FEBS Letters*, *588*(2), 261–8.

Ozcan, G., Ozpolat, B., Coleman, R. L., Sood, A. K., Lopez-Berestein, G. (2015). Preclinical and clinical development of siRNA-based therapeutics. *Advanced Drug Delivery Reviews*.

Sioud, M. (2015). RNA interference: mechanisms, technical challenges, and therapeutic opportunities. *Methods in Molecular Biology (Clifton, N. J.)*, *1218*, 1–15.

Berens, C., Groher, F., Suess, B. (2015). RNA aptamers as genetic control devices: The potential of riboswitches as synthetic elements for regulating gene expression. *Biotechnology Journal*, 10(2), 246–57.

Lönne, M., Zhu, G., Stahl, F., Walter, J.-G. (2014). Aptamer-modified nanoparticles as biosensors. *Advances in Biochemical Engineering/biotechnology*, 140, 121–54.

Genetic engineering: general steps 44/45

http://en.wikipedia.org/wiki/Genetic_engineering

Holloway, B. (1993) *Genetic Exchange Processes for Prokaryotes*, in Biotechnology Second, Completely Revised Edition, 47. Vol. 2, Pühler, A. (ed.). VCH Verlagsgesellschaft, ISBN 3-527-28312-9

Stahl, U., Esser, K. (1993) *Genetic Exchange Processes in Lower Eukaryotes*, in Biotechnology Second, Completely Revised Edition, 73. Vol. 2, Pühler, A. (ed.). VCH Verlagsgesellschaft, ISBN 3-527-28312-9

Enzymes for DNA modification 46/49

http://en.wikipedia.org/wiki/Restriction_enzyme

PCR: general methods and applications 50/53

http://en.wikipedia.org/wiki/PCR

VanGuilder, H. D., Vrana, K. E., Freeman, W. M. (2008). Twenty-five years of quantitative PCR for gene expression analysis. *BioTechniques*, *44*(5), 619–26.

DNA: synthesis and size determination 54/55

http://en.wikipedia.org/wiki/DNA_synthesis

DNA sequencing 56/57

http://en.wikipedia.org/wiki/DNA_sequencing

Middendorf, L., Humphrery, P., Narayanan, N., Roemer, S. (2001) *Sequencing Technology*, in Biotechnology 2nd Edition, 193. Vol. 5b, Sensen, C. (ed.). Wiley-VCH, ISBN 3-527-28328-5

Introduction of foreign DNA in cells 58/59

http://en.wikipedia.org/wiki/Gene_cloning

Schwab, H. (1993) *Principles of Genetic Engineering for Escherichia coli*, in Biotechnology Second, Completely Revised Edition, 373. Vol. 2, Pühler, A. (ed.). VCH Verlagsgesellschaft, ISBN 3-527-28312-9

Wohlleben, W., Muth, G., Kalinowski, J. (1993) *Genetic Engineering of Gram-Positive Bacteria*, in Biotechnology Second, Completely Revised Edition, 455. Vol. 2, Pühler, A. (ed.). VCH Verlagsgesellschaft, ISBN 3-527-28312-9

Priefer, U. (1993) *Principles of Genetic Engineering of Gram-negative Bacteria*, in Biotechnology Second, Completely Revised

Edition, 427. Vol. 2, Pühler, A. (ed.). VCH Verlagsgesellschaft, ISBN 3-527-28312-9
Sudbery, P. (1993) *Genetic Engineering of Yeast*, in Biotechnology Second, Completely Revised Edition, 507. Vol. 2, Pühler, A. (ed.). VCH Verlagsgesellschaft, ISBN 3-527-28312-9
Turner, G. (1993) *Genetic Engineering of Filamentous Fungi*, in Biotechnology Second, Completely Revised Edition, 529. Vol. 2, Pühler, A. (ed.). VCH Verlagsgesellschaft, ISBN 3-527-28312-9

Gene expression 60/61

http://en.wikipedia.org/wiki/Gene_expression
Terpe, K. (2006). Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems, Appl Microbiol Biotechnol, 72(2), 211
Ishihama, A. (2010). Prokaryotic genome regulation: multifactor promoters, multitarget regulators and hierarchic networks. *FEMS Microbiology Reviews*, *34*(5), 628–45.
Riethoven, J.-J. M. (2010). Regulatory regions in DNA: promoters, enhancers, silencers, and insulators. *Methods in Molecular Biology (Clifton, N. J.)*, *674*, 33–42.

Gene silencing 62/63

http://en.wikipedia.org/wiki/Gene_silencing
Ousterout, D. G., Kabadi, A. M., Thakore, P. I., Perez-Pinera, P., Brown, M. T., Majoros, W. H., ... Gersbach, C. A. (2014). Correction of Dystrophin Expression in Cells from Duchenne Muscular Dystrophy Patients through Genomic Excision of Exon 51 by Zinc Finger Nucleases. *Molecular Therapy: The Journal of the American Society of Gene Therapy*.
Jakočiūnas, T., Bonde, I., Herrgård, M., Harrison, S. J., Kristensen, M., Pedersen, L. E., ... Keasling, J. D. (2015). Multiplex metabolic pathway engineering using CRISPR/Cas9 in Saccharomyces cerevisiae. *Metabolic Engineering*.
Dow, L. E., Fisher, J., O'Rourke, K. P., Muley, A., Kastenhuber, E. R., Livshits, G., ... Lowe, S. W. (2015). Inducible in vivo genome editing with CRISPR-Cas9. *Nature Biotechnology*.

Epigenetics 64/65

http://en.wikipedia.org/wiki/Epigenetics
Obata, Y., Furusawa, Y., Hase, K. (2015). Epigenetic modifications of the immune system in health and disease. *Immunology and Cell Biology*.
Jirtle, R. L. (2014). The Agouti mouse: a biosensor for environmental epigenomics studies investigating the developmental origins of health and disease. *Epigenomics*, *6*(5), 447–50.

Genomics 66/73

Brown, T. A. (2006) *Genomes 3,* Garland Science, ISBN 0815341385

Metagenomics 74/75

http://en.wikipedia.org/wiki/Metagenomics
http://en.wikipedia.org/wiki/Global_Ocean_Sampling_Expedition
http://www.brain-biotech.de/en
Sharpton, T. J. (2014). An introduction to the analysis of shotgun metagenomic data. *Frontiers in Plant Science*, *5*, 209.
Owen, J. G., Reddy, B. V. B., Ternei, M. A., Charlop-Powers, Z., Calle, P. Y., Kim, J. H., Brady, S. F. (2013). Mapping gene clusters within arrayed metagenomic libraries to expand the structural diversity of biomedically relevant natural products. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(29), 11797–802.
Williamson, S. J., Rusch, D. B., Yooseph, S., Halpern, A. L., Heidelberg, K. B., Glass, J. I., ... Venter, J. C. (2008). The Sorcerer II Global Ocean Sampling Expedition: metagenomic characterization of viruses within aquatic microbial samples. *PloS One*, *3*(1), e1456.

Cell biology 76–85

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P. (2007), Molecular Biology of the Cell 5th ed., Garland Science, ISBN 978-0-81534105-5
Bolsover, S. R., Shephard, E. A., White, H. A., Hyams, J. S. (2011), Cell biology – a short course, 3rd ed., Wiley-Blackwell, ISBN 978-0-470-52699-6

http://en.wikipedia.org/wiki/Cell_biology

Stem cells 78/79

http://en.wikipedia.org/wiki/Stem_cell
http://stemcells.nih.gov/info/basics/pages/basics1.aspx
Kamran Alimoghaddam (2013) *Stem cells Biology in normal life and disease*, InTech Open Access, ISBN 978-953-51-1107-8

Blood cells and immune system 80/81

http://en.wikipedia.org/wiki/Haematopoiesis
http://en.wikipedia.org/wiki/Immune_system
http://en.wikipedia.org/wiki/Cytokine
http://en.wikipedia.org/wiki/Growth_factor
MacPherson, G., Austyn, J. (2012) *Exploring immunology – concepts and evidence*, Wiley-Blackwell, ISBN 978-3-527-32412-5
Murphy, K. (2011) *Janeway's Immunobiology*, 8th ed., Garland Science, ISBN 978-0-815-34243-4

Antibodies 82/83

http://en.wikipedia.org/wiki/Antibody
http://en.wikipedia.org/wiki/Polyclonal_antibodies
Rajewsky, K. (1996), Clonal selection and learning in the antibody system. *Nature* 381, 751.
Janeway, Jr., C. A. (1993), How the immune system recognizes invaders. *Sci Am* 269, 72.
Jones, R. G. A., Martino, A. (2015). Targeted localized use of therapeutic antibodies: a review of non-systemic, topical and oral applications. *Critical Reviews in Biotechnology*, 1–15.

Reporter groups 84/85

http://en.wikipedia.org/wiki/Reporter_gene
http://en.wikipedia.org/wiki/Digoxigenin
http://en.wikipedia.org/wiki/Flow_cytometry#Fluorescence-activated_cell_sorting_.28FACS.29
Rigler, R. (1995), Fluorescence correlations, single molecule detection and large number screening. Applications in biotechnology, *J Biotechnol* 41, 177.
Voss, U., Larrieu, A., Wells, D. M. (2013). From jellyfish to biosensors: the use of fluorescent proteins in plants. *The International Journal of Developmental Biology*, *57*(6–8), 525–33.
Dunlap, P. (2014). Biochemistry and genetics of bacterial bioluminescence. *Advances in Biochemical Engineering/biotechnology*, *144*, 37–64.
Lundin, A. (2014). Optimization of the firefly luciferase reaction for analytical purposes. *Advances in Biochemical Engineering/biotechnology*, *145*, 31–62.
Juskowiak, B. (2011). Nucleic acid-based fluorescent probes and their analytical potential. *Analytical and Bioanalytical Chemistry*, *399*(9), 3157–76.

Bioengineering 86–109

Takors, R. (2014) Kommentierte Formelsammlung Bioverfahrenstechnik, Springer Spektrum, ISBN 9783642419027

Solid state fermentation 86/87

Mitchell, D. A., Berovic, M., Krieger, N. (2002). Overview of solid state bioprocessing. *Biotechnology Annual Review*, *8*, 183–225.
Durand, A., Renaud, R., Almanza, S., Maratray, J., Diez, M., Desgranges, C. (1993). Solid state fermentation reactors: from lab scale to pilot plant. *Biotechnology Advances*, *11*(3), 591–7.
Leifa, F., Pandey, A., Soccol, C. R. (2000). Solid state cultivation–an efficient method to use toxic agro-industrial residues. *Journal of Basic Microbiology*, *40*(3), 187–97.
Nagel, F. J., Tramper, J., Bakker, M. S., Rinzema, A. (2001). Temperature control in a continuously mixed bioreactor for solid-state fermentation. *Biotechnology and Bioengineering*, 72(2), 219–30.
Durand, A. (2003) Bioreactor designs for solid state fermentation. *Biochem. Eng. Journal* 13, 113-125.
Rodrigues, C., Vandenberghe, L. P. de S., Teodoro, J., Pandey, A., Soccol, C. R. (2009). Improvement on citric acid production in solid-state fermentation by Aspergillus niger LPB BC mutant using citric pulp. *Applied Biochemistry and Biotechnology*, *158*(1), 72–87.
Ercan, D., Demirci, A. (2013) Current and future trends for biofilm reactors for fermentation processes, Crit Rev Biotechnol, Aug 6, PMID 23919241

Growing microorganisms 88/89

http://en.wikipedia.org/wiki/Microbiological_culture

http://en.wikipedia.org/wiki/Growth_medium

Stoppok, E., Buchholz, K. (1996) *Sugar-Based Raw Materials for Fermentation Applications*, in Biotechnology 2nd Edition, 5. *Vol. 6*, Roehr, M. (ed.), Wiley-VCH, ISBN 3-527-28316-1

Growth kinetics and product formation 90/91

http://en.wikipedia.org/wiki/Bacterial_growth

Posten, C. H., Cooney, C. L. (1993) *Growth of Microorganisms* in Biotechnology Second, Completely Revised Edition, 111. Vol. 1, Sahm, H. (ed.). VCH Verlagsgesellschaft, ISBN 3-527-28337-4

Fed-batch and continuous fermentation 92/93

http://en.wikipedia.org/wiki/Fed-batch

Shiloach, J., Fass, R. (2005). Growing E. coli to high cell density–a historical perspective on method development. *Biotechnology Advances*, *23*(5), 345–57.

Chang, H. N., Jung, K., Choi, J.-D.-R., Lee, J. C., Woo, H.-C. Multi-stage continuous high cell density culture systems: a review. *Biotechnology Advances*, *32*(2), 514–25.

Fermentation technology 94/95

http://en.wikipedia.org/wiki/Bioreactor

Katzen, R., Tsao, G. T. (2000), A view of the history of biochemical engineering. *Adv Biochem Eng Biotechnol* 70, 77.

Lee, K. H., Hatzimanikatis, V. (2002), Biochemical Engineering. *Curr Opin Biotechnol* 13, 85.

Mann, U. (2006) *Reactor Technology* in Kirk-Othmer Encyclopedia of Chemical Technology, Wiley Online Library.

Weuster-Botz, D. (2005) Parallel reactor systems for bioprocess development, *Adv Biochem Eng Biotechnol* 92:125.

Loeffelholz, C., Kaiser, S. C., Kraume, M., Eibl, R., Eibl, D. (2014) Dynamic single-use bioreactors ... Engineering characteristics and scaling up, *Adv Biochem. Eng Biotechnol* 97(9):3787.

Fermentation technology: scale-up 96/97

Sonnleitner, B. (2000), Instrumentation of biotechnological processes, *Adv Biochem Eng Biotechnol* 66, 1

Takors, R. (2011) Scale-up of micdrobial processes: impacts, tools and open questions, *J Biotechnol 160*(1–2):3

Schuler, M. M., Marison, I. W. (2012) Real-time monitoring and control of microbial bioprocesses with focus on the specific growth rate: current state and perspectives, A*ppl Microbiol Biotechnol 94*(6):1469

Cultivation of animal cells 98/99

http://en.wikipedia.org/wiki/Animal_cell_culture

http://en.wikipedia.org/wiki/Cell_culture

http://en.wikipedia.org/wiki/Chinese_hamster_ovary_cell

Bardouille, C. (2001) *Maintenance of Cell Cultures – with Special Emphasis on Eukaryotic Cells*, in Biotechnology 2nd Edition, 27. Vol. 10, Rehm, H. (ed.). Wiley-VCH, ISBN 3-527-28320-X

Gramer, M. J. (2014) Product quality considerations for mammalian cell culture process development and manufacturing, *Adv Biochem Eng Biotechnol 139*:123

Kim, J. Y., Kim, Y. G., Lee, G. M. (2012) CHO cells in biotechnology for production of recombinant proteins: current state and further potential, *Appl Microbiol Biotechnol 93*(3):917

Bioreactors for animal cells 100/101

http://en.wikipedia.org/wiki/Minusheet_perfusion_culture_system

Kretzmer, G. (2002), Industrial processes with animal cells. *Appl Microbiol Biotechnol* 59, 135.

Chu, L., Robinson. D. K. (2001), Industrial choices for protein production by large-scale cell culture. *Curr Opin Biotechnol* 12, 180.

Hu, W. S., Aunins, J. G. (1997), Large-scale mammalian cell culture. *Curr Opin Biotechnol* 8, 148.

Kelley, B., Chiou, T., Rosenberg, M., Wang, D. (1993) *Industrial Animal Cell Culture*, in Biotechnology Second, Completely Revised Edition, 23. Vol. 3, Stephanopoulos, G. (ed.). VCH Verlagsgesellschaft, ISBN 3-527-28313-7

Tscheliessnig, A. L., Konrath, J., Bates, R., Jungbauer, A. (2013). Host cell protein analysis in therapeutic protein bioprocessing – methods and applications. *Biotechnology Journal*, *8*(6), 655–70.

Enzyme and cell reactors 102/103

Sheldon, R. A., van Pelt, S. (2013) Enzyme immobilisation in biocatalysis: why, what and how, *Chem Soc Rev 42*(15):6223

Down-stream processing 104/105

http://en.wikipedia.org/wiki/Downstream_(bioprocess)

http://en.wikipedia.org/wiki/Centrifugation

http://en.wikipedia.org/wiki/Liquid–liquid_extractio

Rudolph, R., Lilie, H., Schwarz, E. (1999) *In vitro Folding of Inclusion Body Proteins on an Industrial Scale*, in Biotechnology Second, Completely Revised Edition, 111. Vol. 5a, Ney, U., Schomburg, D. (ed.). VCH Verlagsgesellschaft, ISBN 3-527-28315-3

Mukhopadhyay, A. (1997), Inclusion bodies and purification of proteins in biologically active forms, *Adv Biochem Eng Biotechnol* 56, 61.

Purification of bioproducts 106/107

http://en.wikipedia.org/wiki/Polyhistidine-tag

http://en.wikipedia.org/wiki/Protein_purification

http://en.wikipedia.org/wiki/Affinity_chromatography

Burgess, R., Thompson, N. (2002), Advances in gentle immunoaffinity chromatography. *Curr Opin Biotechnol* 13, 304.

Imamoglu, S: (2002), Simulated moving bed chromatography (SMB) for application in bioseparation. *Adv Biochem Eng Biotechnol* 76, 211.

Queiroz, J. A., Tomaz, C. T., Cabral, J. M. (2001), Hydrophobic interaction chromatography of proteins. *J Biotechnol* 87, 143.

Rathore, A. S., Agarwal, H., Sharma, A. K., Pathak, M., Muthukumar, S. (2015). Continuous processing for production of biopharmaceuticals. *Preparative Biochemistry & Biotechnology*, *45*(8), 836–49.

Food and food additives 110–137

http://en.wikipedia.org/wiki/Fermented_food

Byong, H. Lee (2014) Fundamentals of food biotechnology, 2nd ed., Wiley-Blackwell, ISBN 978-1-118-38491-6

Alcoholic beverages 110/111

http://en.wikipedia.org/wiki/Wine

Pizarro, F., Vargas, F. A., Agosin, E. (2007). A systems biology perspective of wine fermentations. *Yeast (Chichester, England)*, *24*(11), 977–91.

Butzke, C. E., Singleton, V. L. (2007) Wine in Kirk-Othmer Encyclopedia of Chemical Technology, Wiley Online Library

Bartowsky, E. J., Borneman, A. R. (2011). Genomic variations of Oenococcus oeni strains and the potential to impact on malolactic fermentation and aroma compounds in wine. *Applied Microbiology and Biotechnology*, *92*(3), 441–7.

Borneman, A. R., Schmidt, S. A., Pretorius, I. S. (2013). At the cutting-edge of grape and wine biotechnology. *Trends in Genetics: TIG*, *29*(4), 263–71.

Borneman, A. R., Pretorius, I. S., Chambers, P. J. (2013). Comparative genomics: a revolutionary tool for wine yeast strain development. *Current Opinion in Biotechnology*, *24*(2), 192–9.

Beer 112/113

http://en.wikipedia.org/wiki/Beer

Lewis, M. J. (2003) *Beer and Brewing* in Kirk-Othmer Encyclopedia of Chemical Technology, Wiley Interscience

Esslinger, H. M., Narziss, L., (2009) *Beer*, in Ullmann's Encyclopedia of Industrial Chemistry online, Wiley-VCH

Bokulich, N. A., Bamforth, C. W. (2013) The microbiology of malting and brewing, *Microbiol Mol Biol Rev 77*(2):157

Fermented food 114/115

Eugster, E., Jakob, E., Wechsler, D. (2012) Cheese, processed cheese, and whey, Ullmann's Encyclopedia of Industrial Chemistry online. Wiley-VCH

Montel, M. C., Buchin, S., Mallet, A., Delbes-Paus, C., Vuitton, D. A., Desmasures, N., Berthier, F. (2014) Traditional cheeses: rich and diverse microbiota with associated benefits, *Int J Food Microbiol 177*:136

Bachmann, H., Pronk, J. T., Kleerebezem, M., Teusink, B. (2014). Evolutionary engineering to enhance starter culture performance in food fermentations. *Current Opinion in Biotechnology*, 32C, 1–7.

Van Hijum, S. A. F. T., Vaughan, E. E., Vogel, R. F. (2013). Application of state-of-art sequencing technologies to indigenous food fermentations. *Current Opinion in Biotechnology*, 24(2), 178–86.

Food products and lactic acid fermentation 116/117

http://de.wikipedia.org/wiki/Milchsäuregärung

http://en.wikipedia.org/wiki/Sourdough

http://en.wikipedia.org/wiki/Sauerkraut

Smid, E.J., Lacroix, C. (2013) Microbe-microbe interactions in mixed culture food fermentations, *Curr Opin Biotechnol. 24*(2):148

Gaenzle, M. (2014) Enzymatic and bacterial conversions during sourdough fermentation, *Food Microbiol 37*: 2.

Jung, J. Y., Lee, S. H., Jeon, C. O. (2014). Kimchi microflora: history, current status, and perspectives for industrial kimchi production. *Applied Microbiology and Biotechnology*, *98*(6), 2385–93.

Probiotics and Prebiotics 118/119

http://en.wikipedia.org/wiki/Probiotics

http://en.wikipedia.org/wiki/Prebiotic_(nutrition)

http://en.wikipedia.org/wiki/Health_claims_on_food_labels

Gonzalez, A., Clemente, J. C., Shade, A., Metcalf, J. L., Song, S., Prithiviraj, B., ... Knight, R. (2011). Our microbial selves: what ecology can teach us. *EMBO Reports*, *12*(8), 775–84.

Balakrishnan, M., Floch, M. H. (2012). Prebiotics, probiotics and digestive health. *Current Opinion in Clinical Nutrition and Metabolic Care*, *15*(6), 580–5.

Relman, D. A. (2012). The human microbiome: ecosystem resilience and health. *Nutrition Reviews*, *70 Suppl 1*, S2–9.

Ottman, N., Smidt, H., de Vos, W. M., Belzer, C. (2012). The function of our microbiota: who is out there and what do they do? *Frontiers in Cellular and Infection Microbiology*, *2*, 104.

Brown, J., de Vos, W. M., DiStefano, P. S., Doré, J., Huttenhower, C., Knight, R., ... Turnbaugh, P. (2013). Translating the human microbiome. *Nature Biotechnology*, *31*(4), 304–8.

Shoaie, S., Nielsen, J. (2014) Elucidating the interactions between the human gut microbiota and its host through metabolic modelling, *Front Genet 22*(5):86

Vitetta, L., Briskey, D., Alford, H., Hasll, S., Coulson, S: (2014) Probiotics, prebiotics and the gastrointestinal tract in health and disease, *Inflammopharmacology 22*(3):135

Bakers' yeast and fodder yeasts 120/121

http://en.wikipedia.org/wiki/Baker%27s_yeast

http://en.wikipedia.org/wiki/Quorn

Vaughan, P., Macreadie, I. G. (2007) Yeasts, Kirk Othmer Encyclopedia of Chemical Technology.

Single cell protein, single cell oil 122/123

http://en.wikipedia.org/wiki/Single_cell_protein

Litchfield, J. (1994) *Nonconventional Foods*, in Kirk-Othmer Encyclopedia of Chemical Technology 4th Edition 11, 871. Wiley Online Library, ISBN 0-471-52680-0

Babel, W., Pöhland, H.-D., Soyez, K. (1993) *Single Cell Proteins*, in Ullmann's Encyclopedia of Industrial Chemistry 5th edition A24, 165. Wiley-VCH, ISBN 3-527-20124-6

Scrimshaw, N., Murray, E. (1995) *Nutritional Value and Safety of "Single Cell Protein"*, in Biotechnology 2nd Edition, 221. Vol. 9, Reed, G., Nagodawithana, T. (ed.). Wiley-VCH, ISBN 3-527-28319-6

Ratledge, C. (1997) *Microbial Lipids*, in Biotechnology 2nd Edition, 133. Vol. 7, Döhren, H. v., Kleinkauf, H. (ed.). Wiley-VCH, ISBN 3-527-28310-2

Amino acids 124/125

http://en.wikipedia.org/wiki/Amino_acids
Araki, K., Ozeki, T. (2003) *Amino Acids,* in Kirk-Othmer Encyclopedia of Chemical Technology Wiley Online Library,
Drauz, K. H., Grayson, I., Kleemann, A., Krimmer, H. P., Leuchtenberger, W., Weckbauer, C. (2007) *Amino Acids* in Ullmann's Encyclopedia of Industrial Chemistry online,
Becker, J., Wittmann, C. (2012). Systems and synthetic metabolic engineering for amino acid production – the heartbeat of industrial strain development. *Current Opinion in Biotechnology*, *23*(5), 718–26.
Wendisch, V. F. (2014) Microbial production of amino acids and derived chemicals: synthetic biology approaches to strain development, *Curr Opin Biotechnol. 9*, 30, C:51

L-Glutamic acid 126/127

http://en.wikipedia.org/wiki/Glutamic_acid
http://en.wikipedia.org/wiki/Umami
Kawakita, T. (1992) *L-Monosodium Glutamate*, in Kirk-Othmer Encyclopedia of Chemical Technology 4th Edition 2, 571. Wiley Online Library, ISBN 0–471–52669-X
Peters-Wendisch, P. G., Schiel, B., Wendisch, V. F., Katsoulidis, E., Mockel, B., Sahm, H., Eikmanns, B. J. (2001) Pyruvate carboxylase is a major bottleneck for glutamate and lysine production by Corynebacterium glutamicum. *J Mol Microbiol Biotechnol. 3*, 295.

Methionine, L-lysine, L-threonine 128/129

http://en.wikipedia.org/wiki/Methionine
http://en.wikipedia.org/wiki/Lysine
http://en.wikipedia.org/wiki/Threonine
Pfefferle, W., Moeckel, B., Bathe, B., Marx, A. (2003) Biotechnological manufacture of lysine, *Adv. Biochem. Eng Biotechnol 79*:59.

Aspartame, L-Phenylalanine, L-aspartic acid 130/131

http://en.wikipedia.org/wiki/Aspartame
http://en.wikipedia.org/wiki/Advantame
http://en.wikipedia.org/wiki/Phenylalanine
http://en.wikipedia.org/wiki/Sweetener
http://en.wikipedia.org/wiki/Aspartic_acid
Sato, T., Tosa, T. (1993) Production of L-aspartic acid (1993) *Bioprocess Technol. 16*, 15.
Behrens, M., Meyerhof, W., Hellfritsch, C., Hofmann, T. (2011). Sweet and umami taste: natural products, their chemosensory targets, and beyond. *Angewandte Chemie (International Ed. in English)*, *50*(10), 2220–42.

Amino acids via enzymatic transformation 132/133

http://en.wikipedia.org/wiki/Membrane_reactor
Griengl, H., Schwab, H., Fechter, M. (2000), The synthesis of chiral cyanohydrins by oxynitrilases, *Trends Biotechnol* 18, 252.
Bommarius, A., Schwarm, M., Drauz, K. (2001), Comparison of Different Chemoenzymatic Process Routes to Enantiomerically Pure Amino Acids, *Chimia* 55, 50.

Vitamins 133/134

http://en.wikipedia.org/wiki/Vitamins
http://en.wikipedia.org/wiki/Ascorbic_acid
http://en.wikipedia.org/wiki/Riboflavin
http://en.wikipedia.org/wiki/Vitamin_b12
Moine, G., Hohmann, H. P., Kurth, R., Paust, J., Haehnleink, W., Pauling, H., Weimann, B. J., Kaesler, B. (2011) *B Vitamins*, in Ullmann's Encyclopedia of Industrial Chemistry, Wiley Online Library,
Oster, B., Fechtel, U. (2011) *Vitamin C*, in Ullmann's Encyclopedia of Industrial Chemistry, Wiley Online Library,
Eggersdorfer, M., Laudert, D., Létinois, U., McClymont, T., Medlock, J., Netscher, T., Bonrath, W. (2012) One hundred years of vitamins – a success story of the natural sciences, *Angew. Chem. Int Ed Engl. 51*(52):12960

Nucleosides und nucleotides 135/136

http://en.wikipedia.org/wiki/Inosinic_acid
http://en.wikipedia.org/wiki/Umami
Kuninaka, A. (1996) *Nucleotides and Related Compounds*, in Biotechnology 2nd Edition, 561. Vol. 6, Roehr, M. (ed.). Wiley-VCH, ISBN 3-527-28316-1

Asano, Y. (2002). Overview of screening for new microbial catalysts and their uses in organic synthesis–selection and optimization of biocatalysts. *Journal of Biotechnology*, *94*(1), 65–72.

Ethanol 138/139

http://en.wikipedia.org/wiki/Ethanol
http://en.wikipedia.org/wiki/Bioethanol#Bioalcohols
Amorim, H. V, Lopes, M. L., de Castro Oliveira, J. V., Buckeridge, M. S., Goldman, G. H. (2011). Scientific challenges of bioethanol production in Brazil. *Applied Microbiology and Biotechnology*, *91*(5), 1267–75.
Geddes, C. C., Nieves, I. U., Ingram, L. O. (2011). Advances in ethanol production. *Current Opinion in Biotechnology*, *22*(3), 312–319.
Della-Bianca, B. E., Basso, T. O., Stambuk, B. U., Basso, L. C., Gombert, A. K. (2013). What do we know about the yeast strains from the Brazilian fuel ethanol industry? *Applied Microbiology and Biotechnology*, *97*(3), 979–91.
Harris, P. V, Xu, F., Kreel, N. E., Kang, C., Fukuyama, S. (2014). New enzyme insights drive advances in commercial ethanol production. *Current Opinion in Chemical Biology*, 19, 162–70.

1-Butanol, acetone 140/141

http://en.wikipedia.org/wiki/Butanol_fuel
Gu, Y., Jiang, Y., Yang, S., Jiang, W. (2014). Utilization of economical substrate-derived carbohydrates by solventogenic clostridia: pathway dissection, regulation and engineering. *Current Opinion in Biotechnology*, *29C*, 124–131.
Tracy, B. P. (2012). Improving butanol fermentation to enter the advanced biofuel market. *mBio*, *3*(6).
Liu, D., Chen, Y., Ding, F.-Y., Zhao, T., Wu, J.-L., Guo, T., ... Ying, H.-J. (2014). Biobutanol production in a Clostridium acetobutylicum biofilm reactor integrated with simultaneous product recovery by adsorption. *Biotechnology for Biofuels*, *7*(1), 5.
Schiel-Bengelsdorf, B., Montoya, J., Linder, S., Dürre, P. Butanol fermentation. *Environmental Technology*, *34*(13–16), 1691–710.
Green, E. M. (2011). Fermentative production of butanol – the industrial perspective. *Current Opinion in Biotechnology*, *22*(3), 337–43.
Dong, H., Tao, W., Dai, Z., Yang, L., Gong, F., Zhang, Y., Li, Y. (2012). Biobutanol. *Advances in Biochemical Engineering/biotechnology*, *128*, 85–100.

Higher alcohols and alkenes 142/143

http://en.wikipedia.org/wiki/1,3-Propanediol
Nakamura, C. E., Whited, G. M. (2003). Metabolic engineering for the microbial production of 1,3-propanediol. *Current Opinion in Biotechnology*, *14*(5), 454–9.
Zhang, F., Rodriguez, S., Keasling, J. D. (2011). Metabolic engineering of microbial pathways for advanced biofuels production. *Current Opinion in Biotechnology*, *22*(6), 775–83.
Yang, J., Xian, M., Su, S., Zhao, G., Nie, Q., Jiang, X., ... Liu, W. (2012). Enhancing production of bio-isoprene using hybrid MVA pathway and isoprene synthase in E. coli. *PloS One*, *7*(4), e33509.
Tomás, R. A. F., Bordado, J. C. M., Gomes, J. F. P. (2013). p-Xylene oxidation to terephthalic acid: a literature review oriented toward process optimization and development. *Chemical Reviews*, *113*(10), 7421–69.
Buijs, N. A., Siewers, V., Nielsen, J. (2013). Advanced biofuel production by the yeast Saccharomyces cerevisiae. *Current Opinion in Chemical Biology*, 17(3), 480–8.
Chen, Y., Zhou, Y. J., Siewers, V., Nielsen, J. (2014). Enabling Technologies to Advance Microbial Isoprenoid Production. *Advances in Biochemical Engineering/biotechnology*.
Runguphan, W., Keasling, J. D. (2014). Metabolic engineering of Saccharomyces cerevisiae for production of fatty acid-derived biofuels and chemicals. *Metabolic Engineering*, *21*, 103–13.
Jensen, M. K., Keasling, J. D. (2014). Recent applications of synthetic biology tools for yeast metabolic engineering. *FEMS Yeast Research*.

Acetic acid/vinegar 144/145

http://en.wikipedia.org/wiki/Vinegar

Raspor, P., Goranovic, D. (2008) Biotechnological applications of acetic acid bacteria, *Crit Rev Biotechnol. 28*(2):101.

Ebner, H., Follmann, H., Sellmer, S. (1995) *Vinegar*, in Biotechnology 2nd Edition, 579. Vol. 9, Reed, G., Nagodawithana, T. (ed.). Wiley-VCH, ISBN 3-527-28319-6

Ebner, E., Sellmer, S., Follmann, H. (1996) *Acetic Acid*, in Biotechnology 2nd Edition, 381. Vol. 6, Roehr, M. (ed.). Wiley-VCH, ISBN 3-527-28316-1

Ebner, H., Follmann, H., Sellmer, S. (1996) *Vinegar*, in Ullmann's Encyclopedia of Industrial Chemistry 5th edition A27, 403. Wiley-VCH, ISBN 3-527-20127-0

Citric acid 146/147

http://en.wikipedia.org/wiki/Citric_acid

Roehr, M., Kubicek, C., Kominec, J. (1996) *Citric Acid*, in Biotechnology 2nd Edition, 307. Vol. 6, Roehr, M. (ed.). Wiley-VCH, ISBN 3-527-28316-1

Karaffa, L., Kubicek, C. P. (2003). Aspergillus niger citric acid accumulation: do we understand this well working black box? *Applied Microbiology and Biotechnology*, *61*(3), 189–96.

Huang, J., Huang, L., Lin, J., Xu, Z., Cen, P. (2010). *Biotechnology in China II*. (G. T. Tsao, P. Ouyang, & J. Chen, Eds.)*Advances in biochemical engineering/biotechnology* (Vol. 122, pp. 43–71). Springer

Max, B., Salgado, J. M., Rodríguez, N., Cortés, S., Converti, A., Domínguez, J. M. (2010). Biotechnological production of citric acid. *Brazilian Journal of Microbiology: [publication of the Brazilian Society for Microbiology]*, *41*(4), 862–75.

Lactic acid and 3-hydroxypropionic acid 148/149

http://en.wikipedia.org/wiki/Lactic_acid

http://en.wikipedia.org/wiki/3-Hydroxypropionic_acid

Kascak, J., Kominek, J., Roehr, M. (1996) *Lactic Acid*, in Biotechnology 2nd Edition, 293. Vol. 6, Roehr, M. (ed.). Wiley-VCH, ISBN 3-527-28316-1

Jiang, X., Meng, X., Xian, M. (2009). Biosynthetic pathways for 3-hydroxypropionic acid production. *Applied Microbiology and Biotechnology*, *82*(6), 995–1003.

Mimitsuka, T., Na, K., Morita, K., Sawai, H., Minegishi, S., Henmi, M., ... Yonehara, T. (2012). A membrane-integrated fermentation reactor system: its effects in reducing the amount of sub-raw materials for D-lactic acid continuous fermentation by Sporolactobacillus laevolacticus. *Bioscience, Biotechnology, and Biochemistry*, *76*(1), 67–72.

Abdel-Rahman, M. A., Tashiro, Y., Sonomoto, K. (2013). Recent advances in lactic acid production by microbial fermentation processes. *Biotechnology Advances*, *31*(6), 877–902.

Kumar, V., Ashok, S., Park, S. (2013). Recent advances in biological production of 3-hydroxypropionic acid. *Biotechnology Advances*, *31*(6), 945–61.

Gluconic acid and "green" sugar chemicals 150/151

http://en.wikipedia.org/wiki/3-Hydroxypropionic_acid

http://en.wikipedia.org/wiki/Gluconic_acid

http://en.wikipedia.org/wiki/Isosorbide

Roehr, M., Kubicek, C., Kominek, J. (1996) *Gluconic Acid*, in Biotechnology 2nd Edition, 347. Vol. 6, Roehr, M. (ed.). Wiley-VCH, ISBN 3-527-28316-1

Singh, O. V., Kumar, R. (2007). Biotechnological production of gluconic acid: future implications. *Applied Microbiology and Biotechnology*, *75*(4), 713–22.

Van Putten, R.-J., van der Waal, J. C., de Jong, E., Rasrendra, C. B., Heeres, H. J., de Vries, J. G. (2013). Hydroxymethylfurfural, a versatile platform chemical made from renewable resources. *Chemical Reviews*, *113*(3), 1499–597.

http://www.m-kagaku.co.jp/english/products/business/polymer/sustainable/details/1194667_3255.html

Dicarboxylic acids 152/153

http://en.wikipedia.org/wiki/Dicarboxylic_acid

http://en.wikipedia.org/wiki/Succinic_acid

http://en.wikipedia.org/wiki/Itaconic_acid

Andersson, C., Petrova, E., Berglund, K., Rova, U. (2010). Maintaining high anaerobic succinic acid productivity by product removal.

Bioprocess and Biosystems Engineering, *33*(6), 711–8.
Otero, J. M., Cimini, D., Patil, K. R., Poulsen, S. G., Olsson, L., Nielsen, J. (2013). Industrial systems biology of Saccharomyces cerevisiae enables novel succinic acid cell factory. *PloS One*, *8*(1), e54144.
Balzer, G. J., Thakker, C., Bennett, G. N., San, K.-Y. (2013). Metabolic engineering of Escherichia coli to minimize byproduct formate and improving succinate productivity through increasing NADH availability by heterologous expression of NAD(+)-dependent formate dehydrogenase. *Metabolic Engineering*, *20*, 1–8.
Polen, T., Spelberg, M., Bott, M. (2013). Toward biotechnological production of adipic acid and precursors from biorenewables. *Journal of Biotechnology*, *167*(2), 75–84.
Steiger, M. G., Blumhoff, M. L., Mattanovich, D., Sauer, M. (2013). Biochemistry of microbial itaconic acid production. *Frontiers in Microbiology*, *4*, 23.
Alonso, S., Rendueles, M., Díaz, M. (2014). Microbial production of specialty organic acids from renewable and waste materials. *Critical Reviews in Biotechnology*.
Bao, H., Liu, R., Liang, L., Jiang, Y., Jiang, M., Ma, J., ... Ouyang, P. (2014). Succinic acid production from hemicellulose hydrolysate by an Escherichia coli mutant obtained by atmospheric and room temperature plasma and adaptive evolution. *Enzyme and Microbial Technology*, *66*, 10–5.

Biopolymers: Polyesters 154/155

http://en.wikipedia.org/wiki/Polylactic_acid
http://en.wikipedia.org/wiki/Polyhydroxybutyrate
http://en.wikipedia.org/wiki/PHBV
Hazer, B., Steinbüchel, A. (2007). Increased diversification of polyhydroxyalkanoates by modification reactions for industrial and medical applications. *Applied Microbiology and Biotechnology*, *74*(1), 1–12.
Matsumoto, K., Taguchi, S. (2013). Biosynthetic polyesters consisting of 2-hydroxyalkanoic acids: current challenges and unresolved questions. *Applied Microbiology and Biotechnology*, *97*(18), 8011–21.
Tomás, R. A. F., Bordado, J. C. M., Gomes, J. F. P. (2013). p-Xylene oxidation to terephthalic acid: a literature review oriented toward process optimization and development. *Chemical Reviews*, *113*(10), 7421–69.
Meng, D.-C., Shen, R., Yao, H., Chen, J.-C., Wu, Q., Chen, G.-Q. (2014). Engineering the diversity of polyesters. *Current Opinion in Biotechnology*, *29*, 24–33.

Biopolymers: Polyamides 156/157

http://en.wikipedia.org/wiki/Polyamide
http://en.wikipedia.org/wiki/Nylon-12
http://en.wikipedia.org/wiki/Spider_silk
Salerno, A. J., Goldberg, I. (1993). Cloning, expression, and characterization of a synthetic analog to the bioadhesive precursor protein of the sea mussel Mytilus edulis. *Applied Microbiology and Biotechnology*, *39*(2), 221–6.
Carrington, E. (2008). Along the silk road, spiders make way for mussels. *Trends in Biotechnology*, *26*(2), 55–7.
Schacht, K., Scheibel, T. (2014). Processing of recombinant spider silk proteins into tailor-made materials for biomaterials applications. *Current Opinion in Biotechnology*, *29C*, 62–69.

Microbial polysaccharides 158/159

http://en.wikipedia.org/wiki/Xanthan
http://en.wikipedia.org/wiki/Dextran
Sutherland, I. (1996) *Extracellular Polysaccharides*, in Biotechnology 2nd Edition, 613. Vol. 6, Roehr, M. (ed.). Wiley-VCH, ISBN 3-527-28316-1
Brown, S. H., Pummill, P. E. (2008). Recombinant production of hyaluronic acid. *Current Pharmaceutical Biotechnology*, *9*(4), 239–41.
Leemhuis, H., Pijning, T., Dobruchowska, J. M., van Leeuwen, S. S., Kralj, S., Dijkstra, B. W., Dijkhuizen, L. (2013). Glucansucrases: three-dimensional structures, reactions, mechanism, α-glucan analysis and their implications in biotechnology and food applications. *Journal of Biotechnology*, *163*(2), 250–72.
Seviour, R. J., McNeil, B., Fazenda, M. L., Harvey, L. M. (2011). Operating bioreactors for microbial exopolysaccharide production. *Critical Reviews in Biotechnology*, *31*(2), 170–85.

Biosurfactants 160/161

http://en.wikipedia.org/wiki/Sophorolipid#Biosurfactants

http://en.wikipedia.org/wiki/Surfactin

Ron, E. Z., Rosenberg, E. (2002). Biosurfactants and oil bioremediation. *Current Opinion in Biotechnology*, *13*(3), 249–52.

Morita, T., Fukuoka, T., Konishi, M., Imura, T., Yamamoto, S., Kitagawa, M., ... Kitamoto, D. (2009). Production of a novel glycolipid biosurfactant, mannosylmannitol lipid, by Pseudozyma parantarctica and its interfacial properties. *Applied Microbiology and Biotechnology*, *83*(6), 1017–25.

Walter, V., Syldatk, C., Hausmann, R. (2010). Screening concepts for the isolation of biosurfactant producing microorganisms. *Advances in Experimental Medicine and Biology*, *672*, 1–13.

Roelants, S. L. K. W., Saerens, K. M. J., Derycke, T., Li, B., Lin, Y.-C., Van de Peer, Y., ... Soetaert, W. (2013). Candida bombicola as a platform organism for the production of tailor-made biomolecules. *Biotechnology and Bioengineering*, *110*(9), 2494–503.

Amani, H., Müller, M. M., Syldatk, C., Hausmann, R. (2013). Production of microbial rhamnolipid by Pseudomonas aeruginosa MM1011 for ex situ enhanced oil recovery. *Applied Biochemistry and Biotechnology*, *170*(5), 1080–93.

Van Bogaert, I. N. A., Holvoet, K., Roelants, S. L. K. W., Li, B., Lin, Y.-C., Van de Peer, Y., Soetaert, W. (2013). The biosynthetic gene cluster for sophorolipids: a biotechnological interesting biosurfactant produced by Starmerella bombicola. *Molecular Microbiology*, *88*(3), 501–9.

Roelants, S. L. K. W., De Maeseneire, S. L., Ciesielska, K., Van Bogaert, I. N. A., Soetaert, W. (2014). Biosurfactant gene clusters in eukaryotes: regulation and biotechnological potential. *Applied Microbiology and Biotechnology*, *98*(8), 3449–61.

Fatty acids and triglycerides 162/163

http://en.wikipedia.org/wiki/Fatty_acid

http://en.wikipedia.org/wiki/Biodiesel

http://en.wikipedia.org/wiki/Cocoa_butter

Gupta, R., Rathi, P., Bradoo, S. (2003). Lipase mediated upgradation of dietary fats and oils. *Critical Reviews in Food Science and Nutrition*, 43(6), 635–44.

Dillschneider, R., Steinweg, C., Rosello-Sastre, R., Posten, C. (2013). Biofuels from microalgae: photoconversion efficiency during lipid accumulation. *Bioresource Technology*, 142, 647–54.

Sakuradani, E., Ando, A., Shimizu, S., Ogawa, J. (2013). Metabolic engineering for the production of polyunsaturated fatty acids by oleaginous fungus Mortierella alpina 1S-4. *Journal of Bioscience and Bioengineering*, 116(4), 417–22.

Gupta, R., Rathi, P., Bradoo, S. (2003). Lipase mediated upgradation of dietary fats and oils. *Critical Reviews in Food Science and Nutrition*, *43*(6), 635–44.

Dillschneider, R., Steinweg, C., Rosello-Sastre, R., Posten, C. (2013). Biofuels from microalgae: photoconversion efficiency during lipid accumulation. *Bioresource Technology*, *142*, 647–54.

Sakuradani, E., Ando, A., Shimizu, S., Ogawa, J. (2013). Metabolic engineering for the production of polyunsaturated fatty acids by oleaginous fungus Mortierella alpina 1S-4. *Journal of Bioscience and Bioengineering*, *116*(4), 417–22.

Poppe, J. K., Fernandez-Lafuente, R., Rodrigues, R. C., Ayub, M. A. Z. (2015). Enzymatic reactors for biodiesel synthesis: present status and future prospects. *Biotechnology Advances*.

Enzyme Technology 164–199

http://en.wikipedia.org/wiki/Biocatalysis

Buchholz, K., Kasche, V., Bornscheuer, U. T. (2012) *Biocatalysts and Enzyme Technology*, 2nd ed., Wiley-Blackwell, ISBN 978-3-527-32989-2

Biotransformation 164/165

http://en.wikipedia.org/wiki/Biotransformation

http://en.wikipedia.org/wiki/Immobilized_enzyme

Bornscheuer, U. (2000) *Industrial Biotranformations*, in Biotechnology 2nd Edition, 277. Vol. 8b, Kelly, D. (ed.). Wiley-VCH, ISBN 3-527-28324-2

Faber, K. (2000) *Biotransformations in Organic Chemistry*, 4th Edition ed., Springer-Verlag, ISBN 3-540-66334-7

Liese, A., Seelbach, K., Wandrey, C. (2000) *Industrial Biotransformations*, Wiley-VCH, ISBN 3-527-30094-5

Technical enzymes 166/167

http://en.wikipedia.org/wiki/Enzyme_Commission_number

Applied enzyme catalysis 168/169

Ogawa, J., Shimizu, S. (2002). Industrial microbial enzymes: their discovery by screening and use in large-scale production of useful chemicals in Japan. *Current Opinion in Biotechnology*, *13*(4), 367–75.

Jaeger, K.-E., Eggert, T. (2004). Enantioselective biocatalysis optimized by directed evolution. *Current Opinion in Biotechnology*, *15*(4), 305–13.

Panke, S., Wubbolts, M. (2005). Advances in biocatalytic synthesis of pharmaceutical intermediates. *Current Opinion in Chemical Biology*, *9*(2), 188–94.

Ishige, T., Honda, K., Shimizu, S. (2005). Whole organism biocatalysis. *Current Opinion in Chemical Biology*, *9*(2), 174–80.

Kazlauskas, R. J. (2005). Enhancing catalytic promiscuity for biocatalysis. *Current Opinion in Chemical Biology*, *9*(2), 195–201.

Mugford, P. F., Wagner, U. G., Jiang, Y., Faber, K., Kazlauskas, R. J. (2008). Enantiocomplementary enzymes: classification, molecular basis for their enantiopreference, and prospects for mirror-image biotransformations. *Angewandte Chemie (International Ed. in English)*, *47*(46), 8782–93.

Jochens, H., Hesseler, M., Stiba, K., Padhi, S. K., Kazlauskas, R. J., Bornscheuer, U. T. (2011). Protein engineering of α/β-hydrolase fold enzymes. *Chembiochem: A European Journal of Chemical Biology*, *12*(10), 1508–17.

Schober, M., Faber, K. (2013). Inverting hydrolases and their use in enantioconvergent biotransformations. *Trends in Biotechnology*, *31*(8), 468–78.

Köhler, V., Turner, N. J. (2015). Artificial concurrent catalytic processes involving enzymes. *Chemical Communications (Cambridge, England)*, *51*(3), 450–64.

Otte, K. B., Hauer, B. (2015). Enzyme engineering in the context of novel pathways and products. *Current Opinion in Biotechnology*, *35*, 16–22.

Chiral intermediates through biotransformation 170/171

Nestl, B. M., Hammer, S. C., Nebel, B. A., Hauer, B. (2014). New generation of biocatalysts for organic synthesis. *Angewandte Chemie (International Ed. in English)*, *53*(12), 3070–95.

Enzymes as processing aids 172/173

Uhlig, H. (1998) *Industrial Enzymes and their Applications*, John Wiley & Sons, ISBN 0-471-19660-6.

Aehle, A., Misset, O. (1999) *Enzymes for Industrial Applications*, in Biotechnology Second, Completely Revised Edition, 189. Vol. 5a, Ney, U., Schomburg, D. (ed.). VCH Verlagsgesellschaft, ISBN 3-527-28315-3

Kirk, O., Borchert, T. V., Fuglsang, C. C. (2002), Industrial enzyme applications. *Curr Opin in Biotechnol* 13, 345.

Detergent enzymes 174/175

http://en.wikipedia.org/wiki/Laundry_detergent

Ito, S. (1997). Alkaline cellulases from alkaliphilic Bacillus: enzymatic properties, genetics, and application to detergents. *Extremophiles: Life under Extreme Conditions*, *1*(2), 61–6.

Maurer, K.-H. (2004). Detergent proteases. *Current Opinion in Biotechnology*, *15*(4), 330–4.

Despotovic, D., Vojcic, L., Blanusa, M., Maurer, K.-H., Zacharias, M., Bocola, M., ... Schwaneberg, U. (2013). Redirecting catalysis from proteolysis to perhydrolysis in subtilisin Carlsberg. *Journal of Biotechnology*, *167*(3), 279–86.

Niyonzima, F. N., More, S. S. (2014). Detergent-compatible bacterial amylases. *Applied Biochemistry and Biotechnology*, *174*(4), 1215–32.

Vojcic, L., Pitzler, C., Wirtz, G., Jakob, F., Ronny Martinez, Maurer, K.-H., Schwaneberg, U. (2015). Advances in protease engineering for laundry detergents. *New Biotechnology*.

Enzymes for starch degradation 176/177

http://en.wikipedia.org/wiki/Starch
http://en.wikipedia.org/wiki/Amylase
http://en.wikipedia.org/wiki/Starch_gelatinization

Enzymatic starch hydrolysis 178/179

http://en.wikipedia.org/wiki/Starch
http://en.wikipedia.org/wiki/Corn_syrup
http://en.wikipedia.org/wiki/Cyclodextrin

Bertoldo, C., Antranikian, G. (2002) Starch-hydrolyzing enzymes from thermophilic archaea and bacteria. *Curr Opin Chem Biol. 6*, 151.

Tomasik, P., Horton, D. (2012). Enzymatic conversions of starch. *Advances in Carbohydrate Chemistry and Biochemistry*, *68*, 59–436.

Crini, G. (2014). Review: a history of cyclodextrins. *Chemical Reviews*, *114*(21), 10940–75.

Enzymes and sweeteners 180/181

http://en.wikipedia.org/wiki/Sweetener
http://en.wikipedia.org/wiki/High_fructose_corn_syrup
http://en.wikipedia.org/wiki/Invertase

Silveira, M. M., Jonas, R. (2002), The biotechnological production of sorbitol. *Appl Microbiol Biotechnol* 59, 400.

Hebeda, R. (1997) *Syrups*, in Kirk-Othmer Encyclopedia of Chemical Technology 4th Edition 23, 582. Wiley Online Library, ISBN 0-471-52692-4

Mu, W., Li, W., Wang, X., Zhang, T., Jiang, B. (2014). Current studies on sucrose isomerase and biological isomaltulose production using sucrose isomerase. *Applied Microbiology and Biotechnology*, *98*(15), 6569–82.

Enzymes for the degradation of cellulosic biomass 182/183

http://en.wikipedia.org/wiki/Biomass
http://en.wikipedia.org/wiki/Cellulose
http://en.wikipedia.org/wiki/Hemicellulose
http://en.wikipedia.org/wiki/Xylanase
http://en.wikipedia.org/wiki/Cellulase

Thompson, N. (1995) *Hemicellulose* in Kirk-Othmer Encyclopedia of Chemical Technology 4th Edition 13, 54. Wiley Online Library, ISBN 0-471-52682-7

Den Haan, R., van Rensburg, E., Rose, S. H., Görgens, J. F., van Zyl, W. H. (2014). Progress and challenges in the engineering of non-cellulolytic microorganisms for consolidated bioprocessing. *Current Opinion in Biotechnology*, *33C*, 32–38.

Garvey, M., Klose, H., Fischer, R., Lambertz, C., Commandeur, U. (2013). Cellulases for biomass degradation: comparing recombinant cellulase expression platforms. *Trends in Biotechnology*, *31*(10), 581–93.

Hasunuma, T., Okazaki, F., Okai, N., Hara, K. Y., Ishii, J., Kondo, A. (2013). A review of enzymes and microbes for lignocellulosic biorefinery and the possibility of their application to consolidated bioprocessing technology. *Bioresource Technology*, *135*, 513–22.

Hyeon, J. E., Jeon, S. D., Han, S. O. (2013). Cellulosome-based, Clostridium-derived multifunctional enzyme complexes for advanced biotechnology tool development: advances and applications. *Biotechnology Advances*, *31*(6), 936–44.

Enzymatic processes for pulp and paper 184/185

http://en.wikipedia.org/wiki/Pulp_and_paper
http://en.wikipedia.org/wiki/Bleaching_of_wood_pulp

Viikari, L., Tenkanen, M., Suurnäkki, A. (2001) *Biotechnology in the Pulp and Paper Industry*, in Biotechnology 2nd Edition, 523. Vol. 10, Rehm, H. (ed.). Wiley-VCH, ISBN 3-527-28320-X

Polizeli, M. L. T. M., Rizzatti, A. C. S., Monti, R., Terenzi, H. F., Jorge, J. A., Amorim, D. S. (2005). Xylanases from fungi: properties and industrial applications. *Applied Microbiology and Biotechnology*, *67*(5), 577–91.

Gutiérrez, A., del Río, J. C., Martínez, A. T. (2009). Microbial and enzymatic control of pitch in the pulp and paper industry. *Applied Microbiology and Biotechnology*, *82*(6), 1005–18.

Babot, E. D., Rico, A., Rencoret, J., Kalum, L., Lund, H., Romero, J., ... Gutiérrez, A. (2011). Towards industrially-feasible delignification and pitch removal by treating paper pulp with Myceliophthora thermophila laccase

and a phenolic mediator. *Bioresource Technology*, *102*(12), 6717–22.
Torres, C. E., Negro, C., Fuente, E., Blanco, A. (2012). Enzymatic approaches in paper industry for pulp refining and biofilm control. *Applied Microbiology and Biotechnology*, *96*(2), 327–44.

Pectinases 186/187

http://en.wikipedia.org/wiki/Pectin
http://en.wikipedia.org/wiki/Pectinase
Baird, J. (1994) *Gums* in Kirk-Othmer Encyclopedia of Chemical Technology 4th Edition 12, 842. Wiley Online Library, ISBN 0-471-52681-9
Terefe, N. S., Buckow, R., Versteeg, C. (2014). Quality-related enzymes in fruit and vegetable products: effects of novel food processing technologies, part 1: high-pressure processing. *Critical Reviews in Food Science and Nutrition*, *54*(1), 24–63.

Enzymes and milk products 188/189

http://en.wikipedia.org/wiki/Rennet
http://en.wikipedia.org/wiki/Casein
http://en.wikipedia.org/wiki/Cheese
http://en.wikipedia.org/wiki/Galactosemia
http://en.wikipedia.org/wiki/Lactose_intolerance
Hall, C. (1995) *Milk and Milk Products* in Kirk-Othmer Encyclopedia of Chemical Technology 4th Edition 16, 700. Wiley Online Library, ISBN 0-471-52685-1
Kumar, A., Grover, S., Sharma, J., Batish, V. K. (2010). Chymosin and other milk coagulants: sources and biotechnological interventions. *Critical Reviews in Biotechnology*, *30*(4), 243–58.

Enzymes for baking and meat processing 190/191

http://en.wikipedia.org/wiki/Baking
http://en.wikipedia.org/wiki/Tenderizing
Kelly, R. M., Dijkhuizen, L., Leemhuis, H. (2009). Starch and alpha-glucan acting enzymes, modulating their properties by directed evolution. *Journal of Biotechnology*, *140*(3–4), 184–93.
Bekhit, A. A., Hopkins, D. L., Geesink, G., Bekhit, A. A., Franks, P. (2014). Exogenous proteases for meat tenderization. *Critical Reviews in Food Science and Nutrition*, *54*(8), 1012–31.

Other enzymes for food and feed 192/193

http://en.wikipedia.org/wiki/Phytase
http://en.wikipedia.org/wiki/Asparaginase
http://en.wikipedia.org/wiki/Transglutaminase
Kühn, I., Partanen, K. (2012). Phytase improves apparent total tract digestibility of phosphorus and calcium in piglets fed diets with adequate or reduced phosphorus content. *Journal of Animal Science*, *90 Suppl 4*, 194–6.
Itaya, H., Kikuchi, Y. (2008). Secretion of Streptomyces mobaraensis pro-transglutaminase by coryneform bacteria. *Applied Microbiology and Biotechnology*, *78*(4), 621–5.
Hendriksen, H. V, Kornbrust, B. A., Østergaard, P. R., Stringer, M. A. (2009). Evaluating the potential for enzymatic acrylamide mitigation in a range of food products using an asparaginase from Aspergillus oryzae. *Journal of Agricultural and Food Chemistry*, *57*(10), 4168–76.
Humer, E., Schwarz, C., Schedle, K. (2014). Phytate in pig and poultry nutrition. *Journal of Animal Physiology and Animal Nutrition*.

Enzymes for leather and textile treatment 194/195

http://en.wikipedia.org/wiki/Leather#Role_of_enzymes
http://en.wikipedia.org/wiki/Stone_washing
http://en.wikipedia.org/wiki/Desizing
Hamlyn, P. (1995), The Impact of Biotechnology on the Textile Industry *Textiles Magazine* 3, 6.
Heideman, E. (1990) *Leather*, in Ullmann's Encyclopedia of Industrial Chemistry 5th edition A15, 259. Wiley-VCH, ISBN 3-527-20115-7
Li, Q., Yi, L., Marek, P., Iverson, B. L. (2013). Commercial proteases: present and future. *FEBS Letters*, *587*(8), 1155–63.

New procedures towards technical enzymes 196/197

http://en.wikipedia.org/wiki/Protein_engineering

http://en.wikipedia.org/wiki/Lipase
Fischer, M., Pleiss, J. (2003). The Lipase Engineering Database: a navigation and analysis tool for protein families. *Nucleic Acids Research*, 31(1), 319–21.

Protein design and enzyme engineering 198/199

http://en.wikipedia.org/wiki/Directed_evolution
http://en.wikipedia.org/wiki/Protein_design
Reetz, M. T. (2013). Biocatalysis in organic chemistry and biotechnology: past, present, and future. *Journal of the American Chemical Society*, 135(34), 12480–96.
Renata, H., Wang, Z. J., Arnold, F. H. (2015). Expanding the Enzyme Universe: Accessing Non-Natural Reactions by Mechanism-Guided Directed Evolution. *Angewandte Chemie International Edition*.
Nestl, B. M., Hauer, B. (2013). Protein engineering: Chemistry gets the assist. *Nature Chemical Biology*, *9*(8), 470–1.
Turner, N. J. (2009). Directed evolution drives the next generation of biocatalysts. *Nature Chemical Biology*, *5*(8), 567–73.
Ruff, A. J., Dennig, A., Schwaneberg, U. (2013). To get what we aim for–progress in diversity generation methods. *The FEBS Journal*, *280*(13), 2961–78.

Antibiotics 200–221

http://en.wikipedia.org/wiki/Antibiotic

Antibiotics: occurrence and applications 200/201

Hamad, B. (2010). The antibiotics market. *Nature Reviews. Drug Discovery*, *9*(9), 675–6.
Bérdy, J. (2012). Thoughts and facts about antibiotics: Where we are now and where we are heading. *The Journal of Antibiotics*, *65*(8), 441.

Antibiotics: isolation and mechanism of action 202/203

http://www.nature.com/nrd/journal/v12/n5/fig_tab/nrd3975_F1.html
Sanchez, S., Demain, A. L. (2015) Antibiotics – current innovations and future trends, Calster Academic Press, ISBN 978-1-908230-54-6

Antibiotic resistance 204/205

http://en.wikipedia.org/wiki/Antimicrobial_resistance
http://en.wikipedia.org/wiki/Methicillin-resistant_Staphylococcus_aureus
Rennie, R. P. (2012). Current and future challenges in the development of antimicrobial agents. *Handbook of Experimental Pharmacology*, (211), 45–65.
http://www.ecdc.europa.eu/en/healthtopics/antimicrobial_resistance/database/Pages/map_reports.aspx

β-Lactam antibiotics: structure and biosynthesis 206/207

http://en.wikipedia.org/wiki/B-lactam_antibiotic
http://en.wikipedia.org/wiki/Penicillin
http://en.wikipedia.org/wiki/Cephalosporin

β-Lactam antibiotics: manufacture 208/209

Skatrud, P., Schwecke, T., Liempt, H. v., Tobin, M. (1997) *Advances in the Molecular Genetics of β-Lactam Antibiotic Biosynthesis*, in Biotechnology 2nd Edition, 247. Vol. 7, Hv Döhren, Kleinkauf, H. (ed.). Wiley-VCH, ISBN 3-527-28310-2
Wegman, M. A., Janssen, M. H. A., van Rantwijk, F., Sheldon, R. A. (2001), Towards Biocatalytic
Synthesis of β-Lactam Antibiotics. *Adv Synth Catal* 343, 559.
Liu, Y., Gong, G., Zhu, C., Zhu, B., Hu, Y. (2010). Environmentally safe production of 7-ACA by recombinant Acremonium chrysogenum. *Current Microbiology*, *61*(6), 609–14.
Liu, Y., Xie, L., Gong, G., Zhang, W., Zhu, B., Hu, Y. (2014). De novo comparative transcriptome analysis of Acremonium chrysogenum: high-yield and wild-type strains of cephalosporin C producer. *PloS One*, *9*(8), e104542.

Amino acid and peptide antibiotics 210/211

http://en.wikipedia.org/wiki/Lantibiotics
http://en.wikipedia.org/wiki/Cycloserine
http://en.wikipedia.org/wiki/Ciclosporin
http://en.wikipedia.org/wiki/Defensin

Kleinkauf, H., Döhren, H. v. (1997) *Peptide Antibiotics* in Biotechnology 2nd Edition, 277. Vol. 7, Döhren, H. v., Kleinkauf, H. (ed.). Wiley-VCH, ISBN 3-527-28310-2

Glycopeptide, glycolipid, polyether and nucleoside antibiotics 212/213

http://en.wikipedia.org/wiki/Vancomycin
http://en.wikipedia.org/wiki/Methicillin-resistant_Staphylococcus_aureus
http://en.wikipedia.org/wiki/Daptomycin
http://en.wikipedia.org/wiki/Cyclosporin
http://en.wikipedia.org/wiki/Blasticidin_S
Munita, J. M., Murray, B. E., Arias, C. A. (2014). Daptomycin for the treatment of bacteraemia due to vancomycin-resistant enterococci. *International Journal of Antimicrobial Agents*, *44*(5), 387–95.
Peyrani, P., Ramirez, J. (2015). What is the best therapeutic approach to methicillin-resistant Staphylococcus aureus pneumonia? *Current Opinion in Infectious Diseases*.

Aminoglycoside antibiotics 214/215

http://en.wikipedia.org/wiki/Aminoglycoside
http://en.wikipedia.org/wiki/Streptomycin
http://en.wikipedia.org/wiki/Kasugamycin

Tetracyclines, chinolones, and other aromatic antibiotics 216/217

http://en.wikipedia.org/wiki/Tetracyclin
http://en.wikipedia.org/wiki/Fluoroquinolone
http://en.wikipedia.org/wiki/Chloramphenicol
http://en.wikipedia.org/wiki/Doxorubicin
http://en.wikipedia.org/wiki/Ciprofloxacin
Gräfe, U., Dornberger, K., Salz, H. (1997) *Biotechnical Drugs as Antitumor Agents* in Biotechnology 2nd Edition, 641. Vol. 7, Döhren, H. v., Kleinkauf, H. (ed.). Wiley-VCH, ISBN 3-527-28310-2

Macrolide antibiotics 218/219

http://en.wikipedia.org/wiki/Macrolide
http://en.wikipedia.org/wiki/Polyketide
http://en.wikipedia.org/wiki/Amphotericin_B
http://en.wikipedia.org/wiki/Natamycin
http://en.wikipedia.org/wiki/Rifamycin
Kirst, H. (2001) *Antibiotics, Macrolides* in Kirk-Othmer Encyclopedia of Chemical Technology 4th Edition 3, 169. Wiley Online Library, ISBN 0-471-52671-1
Zhong, P., Shortridge, V. (2001) The emerging new generation of antibiotic: ketolides, *Curr Drug Targets Infect Disord. 1*,125.

New procedures for antibiotic synthesis 220/221

Chartrain, M., Salmon, P. M., Robinson, D. K., Buckland, B. C. (2000), Metabolic engineering and directed evolution for the production of pharmaceuticals *Curr Opin in Biotechnology* 11, 209.
Menzella, H. G., Reeves, C. D. (2007). Combinatorial biosynthesis for drug development. *Curr Opin Microbiology*, *10*(3), 238–45.
Wohlleben, W., Mast, Y., Muth, G., Röttgen, M., Stegmann, E., Weber, T. (2012). Synthetic biology of secondary metabolite biosynthesis in actinomycetes: Engineering precursor supply as a way to optimize antibiotic production. *FEBS Letters*, *586*(15), 2171–6.
Baltz, R. H. (2014). Combinatorial biosynthesis of cyclic lipopeptide antibiotics: a model for synthetic biology to accelerate the evolution of secondary metabolite biosynthetic pathways. *ACS Synthetic Biology*, *3*(10), 748–58.
Blin, K., Kazempour, D., Wohlleben, W., Weber, T. (2014). Improved lanthipeptide detection and prediction for antiSMASH. *PloS One*, *9*(2), e89420.

Pharmaceutical and medical biotechnology 222–261

http://en.wikipedia.org/wiki/Biopharmaceutical

Insulin 222/223

http://en.wikipedia.org/wiki/Insulin
Shoelson, S: (1995) *Insulin and Other Antidiabetic Agents*, in Kirk-Othmer Encyclopedia of Chemical Technology 4th Edition 14, 662. Wiley Online Library, ISBN 0-471-52683-5
Tibaldi, J. M. (2014). Evolution of insulin: from human to analog. *The American Journal of Medicine*, *127*(10 Suppl), S25–38.

Growth hormone and other hormones 224/225

http://en.wikipedia.org/wiki/Growth_hormone
http://en.wikipedia.org/wiki/Bovine_somatotropin
http://en.wikipedia.org/wiki/Glucagon
Moreau, D. T. R. (2014). Ecological risk analysis and genetically modified salmon: management in the face of uncertainty. *Annual Review of Animal Biosciences*, *2*, 515–33.
Bernabei, R., Martone, A. M., Ortolani, E., Landi, F., Marzetti, E. (2014). Screening, diagnosis and treatment of osteoporosis: a brief review. *Clinical Cases in Mineral and Bone Metabolism: The Official Journal of the Italian Society of Osteoporosis, Mineral Metabolism, and Skeletal Diseases*, *11*(3), 201–7.
Khan, M., Ouyang, J., Perkins, K., Nair, S., Joseph, F. (2015). Determining predictors of early response to exenatide in patients with type 2 diabetes mellitus. *Journal of Diabetes Research*, *2015*, 162718.

Hemoglobin, serum albumen, lactoferrin 226/227

http://en.wikipedia.org/wiki/Human_serum_albumin
http://en.wikipedia.org/wiki/Lactoferrin
http://en.wikipedia.org/wiki/Hemoglobin
Kobayashi, K. (2006). Summary of recombinant human serum albumin development. *Biologicals: Journal of the International Association of Biological Standardization*, *34*(1), 55–9.

Blood clotting factors 228/229

http://en.wikipedia.org/wiki/Hemophilia
http://en.wikipedia.org/wiki/Factor_VIII
Jiang, R., Monroe, T., McRogers, R., Larson, P. J. (2002) Manufacturing challenges in the commercial production of recombinant coagulation factor VIII, *Haemophilia*. 8 Suppl, 1
Oldenburg, J., Albert, T. (2014). Novel products for haemostasis – current status. *Haemophilia: The Official Journal of the World Federation of Hemophilia*, *20 Suppl 4*, 23–8.

Anticoagulants und thrombolytic agents 230/231

http://en.wikipedia.org/wiki/Tissue_plasminogen_activator
http://en.wikipedia.org/wiki/Urokinase
http://en.wikipedia.org/wiki/Streptokinase
Moser, M., Kohler, B., Schmittner, M., Bode, C. (1998). Recombinant plasminogen activators: a comparative review of the clinical pharmacology and therapeutic use of alteplase and reteplase. *BioDrugs: Clinical Immunotherapeutics, Biopharmaceuticals and Gene Therapy*, *9*(6), 455–63.
Markwardt, F: (2002), Hirudin as alternative anticoagulant – a historical review. *Semin Thromb Hemost* 28, 405.
Werner, R. G., Kopp, K., Schlueter, M. (2007). Glycosylation of therapeutic proteins in different production systems. *Acta Paediatrica (Oslo, Norway: 1992). Supplement*, *96*(455), 17–22.
Longstaff, C., Williams, S., Thelwell, C. (2008). Fibrin binding and the regulation of plasminogen activators during thrombolytic therapy. *Cardiovascular Hematological Agents in Medicinal Chemistry*, *6*(3), 212–23.
Chaudhari, K., Hamad, B., Syed, B. A. (2014). Antithrombotic drugs market. *Nature Reviews. Drug Discovery*, *13*(8), 571–2.

Enzyme inhibitors 232/233

http://en.wikipedia.org/wiki/Trypsin_inhibitor
http://en.wikipedia.org/wiki/Xenical
http://en.wikipedia.org/wiki/Tamiflu
http://en.wikipedia.org/wiki/Acarbose

Interferons 234/235

http://en.wikipedia.org/wiki/Interferon
Wetzel, G: (1999) *Medical Applications of Recombinant Proteins in Humans and Animals*, in Biotechnology Second, Completely Revised Edition, 125. Vol. 5a, Ney, U., Schomburg, D. (ed.). VCH Verlagsgesellschaft, ISBN 3-527-28315-3
Zündorf, I., Dingermann, T. (2014). PEGylation–a well-proven strategy for the improvement of recombinant drugs. *Die Pharmazie*, *69*(5), 323–6.
Huber, A. K., Duncker, P. C., Irani, D. N. (2015). The conundrum of interferon-β non-

responsiveness in relapsing–remitting multiple sclerosis. *Cytokine*.
Webster, D. P., Klenerman, P., Dusheiko, G. M. (2015). Hepatitis C. *Lancet*.

Interleukins and anti-interleukins 236/237

http://en.wikipedia.org/wiki/Interleukin
http://en.wikipedia.org/wiki/Anti-IL-6
McKeage, K. (2014). Ustekinumab: a review of its use in psoriatic arthritis. *Drugs*, *74*(9), 1029–39.
Kim, G. W., Lee, N. R., Pi, R. H., Lim, Y. S., Lee, Y. M., Lee, J. M., ... Chung, S. H. (2015). IL-6 inhibitors for treatment of rheumatoid arthritis: past, present, and future. *Archives of Pharmacal Research*.

Erythropoietin and growth factors 238/239

http://en.wikipedia.org/wiki/Cytokine
http://en.wikipedia.org/wiki/Erythropoietin
http://en.wikipedia.org/wiki/Growth_factor
Hamed, S., Bennett, C. L., Demiot, C., Ullmann, Y., Teot, L., Desmoulière, A. Erythropoietin, a novel repurposed drug: an innovative treatment for wound healing in patients with diabetes mellitus. *Wound Repair and Regeneration: Official Publication of the Wound Healing Society [and] the European Tissue Repair Society*, *22*(1), 23–33.
Momaya, A., Fawal, M., Estes, R. (2015). Performance-Enhancing Substances in Sports: A Review of the Literature. *Sports Medicine (Auckland, N. Z.)*.

Other therapeutic proteins 240/241

http://en.wikipedia.org/wiki/Cystic_fibrosis
http://en.wikipedia.org/wiki/Tumor_necrosis_factors
http://en.wikipedia.org/wiki/Gaucher_disease
Jones, A. P., Wallis, C. (2010). Dornase alfa for cystic fibrosis. *The Cochrane Database of Systematic Reviews*, (3), CD001127.
Schiffmann, R. (2010). Therapeutic approaches for neuronopathic lysosomal storage disorders. *Journal of Inherited Metabolic Disease*, *33*(4), 373–9.
Syed, B. A., Hamad, B. (2014). The cystic fibrosis drug market. *Nature Reviews. Drug Discovery*, *13*(10), 721–2.
Jinesh, S. (2015). Pharmaceutical aspects of anti-inflammatory TNF-blocking drugs. *Inflammopharmacology*.

Monoclonal and catalytic antibodies 242/243

http://en.wikipedia.org/wiki/Monoclonal_antibody
http://en.wikipedia.org/wiki/Abzyme
Blackburn, G., Garcon, A. (2000) *Catalytic Antibodies*, in Biotechnology 2nd Edition, 491. Vol. 8b, Kelly, D. (ed.). Wiley-VCH, ISBN 3-527-28324-2
Werner, R. G. (2004). Economic aspects of commercial manufacture of biopharmaceuticals. *Journal of Biotechnology*, *113*(1–3), 171–82.
Abraham, S., Guo, F., Li, L.-S., Rader, C., Liu, C., Barbas, C. F., ... Sinha, S. C. (2007). Synthesis of the next-generation therapeutic antibodies that combine cell targeting and antibody-catalyzed prodrug activation. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(13), 5584–9.
Chiarella, P. (2011). Production, novel assay development and clinical applications of monoclonal antibodies. *Recent Patents on Anti-Cancer Drug Discovery*, *6*(2), 258–67.
Padiolleau-Lefèvre, S., Ben Naya, R., Shahsavarian, M. A., Friboulet, A., Avalle, B. (2014). Catalytic antibodies and their applications in biotechnology: state of the art. *Biotechnology Letters*, *36*(7), 1369–79.

Recombinant antibodies 244/245

http://en.wikipedia.org/wiki/Monoclonal_antibody
Adair, J. (1999) *Antibody Engineering and Expression*, in Biotechnology Second, Completely Revised Edition, 219. Vol. 5a, Ney, U., Schomburg, D. (ed.). VCH Verlagsgesellschaft, ISBN 3-527-28315-3
Taussig, M. J., Schmidt, R., Cook, E. A., Stoevesandt, O. (2013). Development of proteome-wide binding reagents for research and diagnostics. Proteomics. *Clinical Applications, 7*(11–12), 756–66.

Zhao, A., Tohidkia, M. R., Siegel, D. L., Coukos, G., Omidi, Y. (2014). Phage antibody display libraries: a powerful antibody discovery platform for immunotherapy. *Critical Reviews in Biotechnology*, 1–14.

Nishimiya, D. (2014). Proteins improving recombinant antibody production in mammalian cells. *Applied Microbiology and Biotechnology, 98*(3), 1031–42.

Doerner, A., Rhiel, L., Zielonka, S., Kolmar, H. (2014). Therapeutic antibody engineering by high efficiency cell screening. *FEBS Letters, 588*(2), 278–87.

Therapeutic antibodies 246/247

http://en.wikipedia.org/wiki/Monoclonal_antibody_therapy

http://en.wikipedia.org/wiki/Monoclonal_antibody

http://fabian-weber.eu/wp-content/uploads/2012/11/TOP1201.png

Racher, A., Tong, J., Bonnerjea, J. (1999) *Manufacture of Therapeutic Antibodies* in Biotechnology Second, Completely Revised Edition, 247. Vol. 5a, Ney, U., Schomburg, D. (ed.). VCH Verlagsgesellschaft, ISBN 3-527-28315-3

Vaccines 248/249

http://en.wikipedia.org/wiki/Vaccine

Hsieh, C., Ritchey, M. (1997) *Vaccine Technology*, in Kirk-Othmer Encyclopedia of Chemical Technology 4th Edition 24, 727. Wiley Online Library, ISBN 0-471-52693-2

Robbins, M. J., Jacobson, S. H. (2014). Analytics for vaccine economics and pricing: insights and observations. *Expert Review of Vaccines*, 1–12.

Recombinant vaccines 250/251

http://en.wikipedia.org/wiki/Vaccine

http://en.wikipedia.org/wiki/Hepatitis_B_vaccine

http://en.wikipedia.org/wiki/H5N1_vaccine

Paul, M., van Dolleweerd, C., Drake, P. M. W., Reljic, R., Thangaraj, H., Barbi, T., ... Ma, J. K.-C. (2011). Molecular Pharming: future targets and aspirations. *Human Vaccines*, *7*(3), 375–82.

De Groot, A. S., Einck, L., Moise, L., Chambers, M., Ballantyne, J., Malone, R. W., ... Martin, W. (2013). Making vaccines "on demand": a potential solution for emerging pathogens and biodefense? *Human Vaccines & Immunotherapeutics*, *9*(9), 1877–84.

Geall, A. J., Mandl, C. W., Ulmer, J. B. (2013). RNA: the new revolution in nucleic acid vaccines. *Seminars in Immunology*, *25*(2), 152–9.

Hess, R. D., Weber, F., Watson, K., Schmitt, S. (2012). Regulatory, biosafety and safety challenges for novel cells as substrates for human vaccines. *Vaccine*, *30*(17), 2715–27.

Steroid biotransformations 252/253

http://en.wikipedia.org/wiki/Steroid

Müller, R. (1994) *Steroids* in Ullmann's Encyclopedia of Industrial Chemistry 5th edition A25, 309. Wiley-VCH, ISBN 3-527-20125-4

Duport, C., Spagnoli, R., Degryse, E., Pompon, D. (1998). Self-sufficient biosynthesis of pregnenolone and progesterone in engineered yeast. *Nature Biotechnology*, *16*(2), 186–9.

Bernhardt, R. (2006). Cytochromes P450 as versatile biocatalysts. *Journal of Biotechnology*, *124*(1), 128–45.

Urban, P., Truan, G., Pompon, D. (2014). High-throughput functional screening of steroid substrates with wild-type and chimeric P450 enzymes. *BioMed Research International*, *2014*, 764102.

Enzymes for analysis 254/255

http://en.wikipedia.org/wiki/Enzyme_assay

Usmani, A. (1995) *Medical Diagnostic Reagents*, in Kirk-Othmer Encyclopedia of Chemical Technology 4th Edition 16, 88. Wiley Online Library, ISBN 0-471-52685-1

Kresse, G. (1995) *Analytical Use of Enzymes*, in Biotechnology 2nd Edition, 137. Vol. 9, Reed, G., Nagodawithana, T: (ed.). Wiley-VCH, ISBN 3-527-28319-6

Enzyme tests 256/257

http://en.wikipedia.org/wiki/Enzyme_assay

http://en.wikipedia.org/wiki/Liver_function_tests

Gherson, P., Lanza, H., Elavin, M., Vlastelica, D.(1992) *Automated Instrumenation, Clincial Chemistry*, in Kirk-Othmer Encyclopedia

of Chemical Technology 4th Edition 3, 751. Wiley Online Library, ISBN 0-471-52671-1

Biosensors 258/259

http://en.wikipedia.org/wiki/Biosensor
Bier, F. F., Schumacher, S. (2013). Integration in bioanalysis: technologies for point-of-care testing. *Advances in Biochemical Engineering/biotechnology*, *133*, 1–14.

Immunoassays 260/261

http://en.wikipedia.org/wiki/Immunoassay
http://en.wikipedia.org/wiki/Troponin_T
Miller, J., Niessner, R. (1994) *Enzyme and Immunoassays*, in Ullmann's Encyclopedia of Industrial Chemistry 5th edition B5, 129. Wiley-VCH, ISBN 3-527-20135-1
Chan, C. P., Cheung, Y., Renneberg, R., Seydack, M. (2008). New trends in immunoassays. *Advances in Biochemical Engineering/biotechnology*, *109*, 123–54.

Glycobiology 262/263

http://en.wikipedia.org/wiki/Glycobiology
Aricescu, A. R., Owens, R. J. (2013). Expression of recombinant glycoproteins in mammalian cells: towards an integrative approach to structural biology. *Current Opinion in Structural Biology*, *23*(3), 345–56.
http://www.ncbi.nlm.nih.gov/books/NBK1898/: Mulloy, B., Hart, G. W., Stanley, P. (2009). Structural Analysis of Glycans. Cold Spring Harbor Laboratory Press.
Hossler, P., Khattak, S. F., Li, Z. J. (2009). Optimal and consistent protein glycosylation in mammalian cell culture. *Glycobiology*, *19*(9), 936–49.
Loos, A., Steinkellner, H. (2014). Plant glycobiotechnology on the way to synthetic biology. *Frontiers in Plant Science*, *5*, 523.
Nevalainen, H., Peterson, R. (2014). Making recombinant proteins in filamentous fungi-are we expecting too much? *Frontiers in Microbiology*, *5*, 75.
Hagan, A. K., Wang, M., Liu, L. (2014). Current approaches to glycoprotein analysis. *Protein and Peptide Letters*, *21*(10), 986–99.

Animal breeding 264/265

http://en.wikipedia.org/wiki/Animal_breeding
Gianola, D., Rosa, G. J. M. (2015). One hundred years of statistical developments in animal breeding. *Annual Review of Animal Biosciences*, *3*, 19–56.
Washburn, S. P., Mullen, K. A. E. (2014). Invited review: Genetic considerations for various pasture-based dairy systems. *Journal of Dairy Science*, *97*(10), 5923–38.
Buckley, F., Lopez-Villalobos, N., Heins, B. J. (2014). Crossbreeding: implications for dairy cow fertility and survival. *Animal: An International Journal of Animal Bioscience*, *8 Suppl 1*, 122–33.
Hansen, P. J. (2014). Current and future assisted reproductive technologies for mammalian farm animals. *Advances in Experimental Medicine and Biology*, *752*, 1–22.
Seidel, G. E. (2014). Beef cattle in the year 2050. *Advances in Experimental Medicine and Biology*, *752*, 239–44.

Embryo transfer, cloned animals 266/267

http://en.wikipedia.org/wiki/Embryo_transfer
http://en.wikipedia.org/wiki/Cloning
http://en.wikipedia.org/wiki/Mouse_models_of_human_disease
Wilmut, I., Sullivan, G., Taylor, J. (2009). A decade of progress since the birth of Dolly. *Reproduction, Fertility, and Development*, *21*(1), 95–100.
Graf, A., Krebs, S., Heininen-Brown, M., Zakhartchenko, V., Blum, H., Wolf, E. (2014). Genome activation in bovine embryos: review of the literature and new insights from RNA sequencing experiments. *Animal Reproduction Science*, *149*(1–2), 46–58.
Bó, G. A., Baruselli, P. S. (2014). Synchronization of ovulation and fixed-time artificial insemination in beef cattle. *Animal: An International Journal of Animal Bioscience*, *8 Suppl 1*, 144–50.
McDougall, S., Heuer, C., Morton, J., Brownlie, T. (2014). Use of herd management programmes to improve the reproductive performance of dairy cattle. *Animal: An International Journal of Animal Bioscience*, *8 Suppl 1*, 199–210.
Ponsart, C., Le Bourhis, D., Knijn, H., Fritz, S., Guyader-Joly, C., Otter, T., ... Mullaart, E.

(2013). Reproductive technologies and genomic selection in dairy cattle. *Reproduction, Fertility, and Development*, *26*(1), 12–21.

Wiltbank, M. C., Pursley, J. R. (2014). The cow as an induced ovulator: timed AI after synchronization of ovulation. *Theriogenology*, *81*(1), 170–85.

Gene mapping 268/269

http://en.wikipedia.org/wiki/Gene_mapping

Pérez-Enciso, M., Mercadé, A., Bidanel, J. P., Geldermann, H., Cepica, S., Bartenschlager, H., ... Folch, J. M. (2005). Large-scale, multibreed, multitrait analyses of quantitative trait loci experiments: the case of porcine X chromosome. *Journal of Animal Science*, *83*(10), 2289–96.

Geldermann, H., Cepica, S., Stratil, A., Bartenschlager, H., Preuss, S. (2010). Genome-wide mapping of quantitative trait loci for fatness, fat cell characteristics and fat metabolism in three porcine F2 crosses. *Genetics, Selection, Evolution: GSE*, *42*, 31.

Berry, D. P., Wall, E., Pryce, J. E. (2014). Genetics and genomics of reproductive performance in dairy and beef cattle. *Animal: An International Journal of Animal Bioscience*, *8 Suppl 1*, 105–21.

Egger-Danner, C., Cole, J. B., Pryce, J. E., Gengler, N., Heringstad, B., Bradley, A., Stock, K. F. (2015). Invited review: overview of new traits and phenotyping strategies in dairy cattle with a focus on functional traits. *Animal: An International Journal of Animal Bioscience*, *9*(2), 191–207.

Transgenic animals 270/271

http://en.wikipedia.org/wiki/Mouse_models_of_human_disease

Palmiter, R. D., Brinster, R. L., Hammer, R. E., Trumbauer, M. E., Rosenfeld, M. G., Birnberg, N. C., Evans, R. M. (1982). Dramatic growth of mice that develop from eggs microinjected with metallothionein-growth hormone fusion genes. *Nature*, *300*(5893), 611–5.

Ayadi, A., Birling, M.-C., Bottomley, J., Bussell, J., Fuchs, H., Fray, M., ... Herault, Y. (2012). Mouse large-scale phenotyping initiatives: overview of the European Mouse Disease Clinic (EUMODIC) and of the Wellcome Trust Sanger Institute Mouse Genetics Project. *Mammalian Genome: Official Journal of the International Mammalian Genome Society*, *23*(9–10), 600–10.

Mouse genome informatics: http://www.informatics.jax.org

Gene farming and xenotransplantation 272/273

http://en.wikipedia.org/wiki/Pharming_(genetics)

http://en.wikipedia.org/wiki/Xenotransplantation

Dove, A. (2000). Milking the genome for profit. *Nature Biotechnology*, *18*(10), 1045–8.

Niemann, H., Kues, W. A. (2007). Transgenic farm animals: an update. *Reproduction, Fertility, and Development*, *19*(6), 762–70.

Maksimenko, O. G., Deykin, A. V, Khodarovich, Y. M., Georgiev, P. G. (2013). Use of transgenic animals in biotechnology: prospects and problems. *Acta Naturae*, *5*(1), 33–46.

Cooper, D. K. C., Satyananda, V., Ekser, B., van der Windt, D. J., Hara, H., Ezzelarab, M. B., Schuurman, H.-J. Progress in pig-to-non-human primate transplantation models (1998–2013): a comprehensive review of the literature. *Xenotransplantation*, *21*(5), 397–419.

Michel, S. G., Madariaga, M. L. L., Villani, V., Shanmugarajah, K. (2015). Current progress in xenotransplantation and organ bioengineering. *International Journal of Surgery (London, England)*, *13C*, 239–244.

Plant breeding 274/275

http://en.wikipedia.org/wiki/Plant_breeding

http://faostat.fao.org/site/567/desktopdefault.aspx#ancor

Bohra, A., Pandey, M. K., Jha, U. C., Singh, B., Singh, I. P., Datta, D., ... Varshney, R. K. (2014). Genomics-assisted breeding in four major pulse crops of developing countries: present status and prospects. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *127*(6), 1263–91.

Huang, J., Pray, C., Rozelle, S. (2002). Enhancing the crops to feed the poor. *Nature*, *418*(6898), 678–84.

Palmgren, M. G., Edenbrandt, A. K., Vedel, S. E., Andersen, M. M., Landes, X., Østerberg, J. T., ... Pagh, P. (2014). Are we ready for back-to-nature crop breeding? *Trends in Plant Science*.

Desta, Z. A., Ortiz, R. (2014). Genomic selection: genome-wide prediction in plant improvement. *Trends in Plant Science*, *19*(9), 592–601.

Yamagishi, H., Bhat, S. R. (2014). Cytoplasmic male sterility in Brassicaceae crops. *Breeding Science*, *64*(1), 38–47.

Clevenger, J., Chavarro, C., Pearl, S. A., Ozias-Akins, P., Jackson, S. A. (2015). SNP identification in polyploids: a review, example and recommendations. *Molecular Plant*.

Shi, J., Lai, J. (2015). Patterns of genomic changes with crop domestication and breeding. *Current Opinion in Plant Biology*, *24C*, 47–53.

Cultivation of plant cells 276/277

http://en.wikipedia.org/wiki/Plant_tissue_culture

Grout, B. W. (1999). Meristem-tip culture for propagation and virus elimination. *Methods in Molecular Biology (Clifton, N.J.)*, *111*, 115–25.

Grout, B. W. W. (2007). Cryopreservation of plant cell suspensions. *Methods in Molecular Biology (Clifton, N.J.)*, *368*, 153–61.

Ghanem, M. E., Marrou, H., Sinclair, T. R. (2014). Physiological phenotyping of plants for crop improvement. *Trends in Plant Science*.

Plant tissue culture and plant breeding 278/279

http://en.wikipedia.org/wiki/Plant_tissue_culture

Breuling, M., Alfermann, A. W., Reinhard, E. (1985). Cultivation of cell cultures of Berberis wilsonae in 20-l airlift bioreactors. *Plant Cell Reports*, *4*(4), 220–3.

Kreis, W., Reinhard, E. (1992). 12 beta-Hydroxylation of digitoxin by suspension-cultured Digitalis lanata cells: production of digoxin in 20-litre and 300-litre air-lift bioreactors. *Journal of Biotechnology*, *26*(2–3), 257–73.

Ge, F., Yuan, X., Wang, X., Zhao, B., Wang, Y. (2006). Cell growth and shikonin production of Arnebia euchroma in a periodically submerged airlift bioreactor. *Biotechnology Letters*, *28*(8), 525–9.

Georgiev, M. I., Weber, J. (2014). Bioreactors for plant cells: hardware configuration and internal environment optimization as tools for wider commercialization. *Biotechnology Letters*, *36*(7), 1359–67.

Murthy, H. N., Georgiev, M. I., Kim, Y.-S., Jeong, C.-S., Kim, S.-J., Park, S.-Y., Paek, K.-Y. (2014). Ginsenosides: prospective for sustainable biotechnological production. *Applied Microbiology and Biotechnology*, *98*(14), 6243–54.

Transgenic plants: methods 280/281

http://en.wikipedia.org/wiki/Transgenic_soybean

http://en.wikipedia.org/wiki/Ti_plasmid

http://en.wikipedia.org/wiki/Gene_gun

Vanavichit, A., Tragoonrung, S., Toojinda, T. (2001) *Genomic Mapping and Positional Cloning with Emphasis on Plant Science*, in Biotechnology 2nd Edition, 165. Vol. 5b, Sensen, C. (ed.). Wiley-VCH, ISBN 3-527-28328-5

Bortesi, L., Fischer, R. (2014). The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnology Advances*, *33*(1), 41–52.

Langridge, P., Reynolds, M. P. (2014). Genomic tools to assist breeding for drought tolerance. *Current Opinion in Biotechnology*, *32C*, 130–135.

Hou, H., Atlihan, N., Lu, Z.-X. (2014). New biotechnology enhances the application of cisgenesis in plant breeding. *Frontiers in Plant Science*, *5*, 389.

Szabala, B. M., Osipowski, P., Malepszy, S. (2014). Transgenic crops: the present state and new ways of genetic modification. *Journal of Applied Genetics*, *55*(3), 287–94.

Transgenic plants: resistance 282/283

http://en.wikipedia.org/wiki/Transgenic_plants

http://en.wikipedia.org/wiki/Genetically_modified_crops

Mullins, E. (2014). Engineering for disease resistance: persistent obstacles clouding tangible opportunities. *Pest Management Science*.

Tanaka, Y., Brugliera, F. (2013). Flower colour and cytochromes P450. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *368*(1612), 20120432.

Genetically Modified Pest-Protected Plants: Science and Regulation – PubMed – NCBI. (n. d.). Retrieved February 21, 2015, from http://www.ncbi.nlm.nih.gov/pubmed/25032472

Statistics on Transgenic Plants: http://www.isaaa.org/resources/videos/globalstatusreport2014/default.asp

Transgenic plants: applications 284/285

http://en.wikipedia.org/wiki/Transgenic_plants

http://en.wikipedia.org/wiki/Transgenic_maize

http://en.wikipedia.org/wiki/Transgenic_wheat

http://en.wikipedia.org/wiki/Transgenic_tomato

Dehesh, K., Jones, A., Knutzon, D. S., Voelker, T. A. (1996). Production of high levels of 8:0 and 10:0 fatty acids in transgenic canola by overexpression of Ch FatB2, a thioesterase cDNA from Cuphea hookeriana. *The Plant Journal: For Cell and Molecular Biology*, *9*(2), 167–72.

Giddings, G. (2001). Transgenic plants as protein factories. *Current Opinion in Biotechnology*, *12*(5), 450–4.

Carrari, F., Urbanczyk-Wochniak, E., Willmitzer, L., Fernie, A. R. (2003). Engineering central metabolism in crop species: learning the system. *Metabolic Engineering*, *5*(3), 191–200.

Van Beilen, J. B., Poirier, Y. (2008). Production of renewable polymers from crop plants. *The Plant Journal: For Cell and Molecular Biology*, *54*(4), 684–701.

Storck, T. Amflora: great expectation for GM crops in Europe. Interview by Naglaa Abdallah. *GM Crops*, *1*(3), 109–12.

Osakabe, Y., Osakabe, K. (2014). Genome Editing with Engineered Nucleases in Plants. *Plant & Cell Physiology*.

Porth, I., El-Kassaby, Y. A. (2015). Using Populus as a lignocellulosic feedstock for bioethanol. *Biotechnology Journal*.

Aerobic wastewater treatment 286/287

http://en.wikipedia.org/wiki/Sewage_treatment

Nisipeanu, P. (1999) *Laws, Statutory Orders and Directives on Waste and Wastewater Treatment* in Biotechnology 2nd Edition, 141. Vol. 11a, Winter, J. (ed.). Wiley-VCH, ISBN 3-527-28321-8

Gallert, C., Winter, J. (2000) *Perspectives of Waste, Wastewater, Off-Gas, and Drinking Water Management*, in Biotechnology 2nd Edition, 479. Vol. 11c, Winter, J. (ed.). Wiley-VCH, ISBN 3-527-28336-6

Fritsche, W., Hofrichter, M. (2000) *Aerobic Degradation by Microorganisms*, in Biotechnology 2nd Edition, 145. Vol. 11b, Klein, J. (ed.). Wiley-VCH, ISBN 3-527-28323-4

Hartmann, L. (1999) *Historical Development of Wastewater Treatment Processes*, in Biotechnology 2nd Edition, 5. Vol. 11a, Winter, J. (ed.). Wiley-VCH, ISBN 3-527-28321-8

Gallert, C., Winter, J. (1999) *Bacterial Metabolism in Wastewater Treatment*, in Biotechnology 2nd Edition, 17. Vol. 11a, Winter, J. (ed.). Wiley-VCH, ISBN 3-527-28321-8

Koppe, P., Stozek, A., Neitzel, V. (1999) *Municipal Wastewater and Sewage Sludge*, in Biotechnology 2nd Edition, 161. Vol. 11a, Winter, J. (ed.). Wiley-VCH, ISBN 3-527-28321-8

Rosenwinkel, K., Austermann-Haun, U., Meyer, H. (1999) *Industrial Wastewater Sources and Treatment Strategies*, in Biotechnology 2nd Edition, 191. Vol. 11a, Winter, J. (ed.). Wiley-VCH, ISBN 3-527-28321-8

Weiland, P. (1999) *Agricultural Waste and Wastewater Sources and Management*, in Biotechnology 2nd Edition, 217. Vol. 11a, Winter, J. (ed.). Wiley-VCH, ISBN 3-527-28321-8

Riedel, K., Kunze, G., König, A. (2002). Microbial sensors on a respiratory basis for wastewater monitoring. *Advances in Biochemical Engineering/biotechnology*, *75*, 81–118.

Kjelleberg, S. (2002). Environmental biotechnology. *Current Opinion in Biotechnology*, *13*(3), 199–203.

Anaerobic waste water and sludge treatment 288/289

http://en.wikipedia.org/wiki/Sewage_treatment

http://en.wikipedia.org/wiki/Biogas

Gallert, C., Winter, J. (2002) Solid and liquid residues as raw materials for biotechnology, *Naturwissenschaften*. 89, 483

Sekiguchi, Y., Kamagata, Y., Harada, H. (2001), Recent advances in methane fermentation technology *Curr Opin Biotechnol* 12, 277.

Schink, B. (2000) *Principles of Anaerobic Degradation of Organic Compounds*, in Biotechnology 2nd Edition, 169. Vol. 11b, Klein, J. (ed.). Wiley-VCH, ISBN 3-527-28323-4

McInerney, M. (1999) *Anaerobic Metabolism and its Regulation*, in Biotechnology 2nd Edition, 455. Vol. 11a, Winter, J. (ed.). Wiley-VCH, ISBN 3-527-28321-8

Märkl, H. (1999) *Modeling of Biogas Reactors*, in Biotechnology 2nd Edition, 527. Vol. 11a, Winter, J. (ed.). Wiley-VCH, ISBN 3-527-28321-8

Cecchi, F., & Cavinato, C. (2015). Anaerobic digestion of bio-waste: A mini-review focusing on territorial and environmental aspects. *Waste Management & Research: The Journal of the International Solid Wastes and Public Cleansing Association, ISWA*.

Biological waste gas treatment 290/291

Engesser, K., Plaggemeier, T. (2000) *Microbiological Aspects of Biological Waste Gas Purification*, in Biotechnology 2nd Edition, 275. Vol. 11c, Winter, J. (ed.). Wiley-VCH, ISBN 3-527-28336-6

Fischer, K. (2000) *Biofilters*, in Biotechnology 2nd Edition, 321. Vol. 11c, Winter, J. (ed.). Wiley-VCH, ISBN 3-527-28336-6

Chitwood, D., Devinny, J. (2000) *Commercial Applications of Biological Waste Gas Purification,* in Biotechnology 2nd Edition, 357. Vol. 11c, Winter, J. (ed.). Wiley-VCH, ISBN 3-527-28336-6

Plaggemeier, T., Lämmerzahl, O. (2000) *Treatment of Waste Gas Pollutants in Trickling Filters*, in Biotechnology 2nd Edition, 333. Vol. 11c, Winter, J. (ed.). Wiley-VCH, ISBN 3-527-28336-6

Reiser, M. (2000) *Waste Gas Treatment: Membrane Processes and Alternative Techniques*, in Biotechnology 2nd Edition, 345. Vol. 11c, Winter, J. (ed.). Wiley-VCH, ISBN 3-527-28336-6

Schippert, E., Chmiel, H. (2000) *Bioscrubbers*, in Biotechnology 2nd Edition, 305. Vol. 11c, Winter, J. (ed.). Wiley-VCH, ISBN 3-527-28336-6

Waweru, M., Herrygers, V., Langenhove, H. V., Verstraete, W: (2000) *Process Engineering of Biological Waste Gas Purification*, in Biotechnology 2nd Edition, 259. Vol. 11c, Winter, J. (ed.). Wiley-VCH, ISBN 3-527-28336-6

Biological treatment of polluted soil 292/293

http://en.wikipedia.org/wiki/Bioremediation

Ulrici, W. (2000) *Contaminated Soil Areas, Different Countries and Contaminants, Monitoring of Contaminants*, in Biotechnology 2nd Edition, 5. Vol. 11b, Klein, J. (ed.). Wiley-VCH, ISBN 3-527-28323-4

Klein, J. (2000) *Possibilities, Limits, and Future Developments of Soil Bioremediation*, in Biotechnology 2nd Edition, 465. Vol. 11b, Klein, J. (ed.). Wiley-VCH, ISBN 3-527-28323-4

Held, T., Dörr, H. (2000) *In situ Remediation*, in Biotechnology 2nd Edition, 349. Vol. 11b, Klein, J. (ed.). Wiley-VCH, ISBN 3-527-28323-4

Unterman, R., DeFlaun, M., Steffan, S.J. (2000) *Acvanced in situ Bioremediation – A Hierarchy of Technology Choices*, in Biotechnology 2nd Edition, 399. Vol. 11b, Klein, J. (ed.). Wiley-VCH, ISBN 3-527-28323-4

Wischnak, C., Müller, R. (2000) *Degradation of Chlorinated Compounds*, in Biotechnology 2nd Edition, 241. Vol. 11b, Klein, J. (ed.). Wiley-VCH, ISBN 3-527-28323-4

Schulz-Berendt, V. (2000) *Bioremediation with Heap Technique*, in Biotechnology 2nd Edition, 319. Vol. 11b, Klein, J. (ed.). Wiley-VCH, ISBN 3-527-28323-4

Blotevogel, K., Gorontzy, T. (2000) *Microbial Degradation of Compounds with Nitro Functions*, in Biotechnology 2nd Edition, 273. Vol. 11b, Klein, J. (ed.). Wiley-VCH, ISBN 3-527-28323-4

Menn, F.-M., Easter, J., Sayleri, G. (2000) *Genetically Engineered Microorganisms and Bioremediation*, in Biotechnology 2nd Edi-

tion, 5. Vol. 11b, Klein, J. (ed.). Wiley-VCH, ISBN 3-527-28323-4

Urgun-Demirtas, M., Stark, B., Pagilla, K. Use of genetically engineered microorganisms (GEMs) for the bioremediation of contaminants. *Critical Reviews in Biotechnology*, *26*(3), 145–64.

Vila, J., Tauler, M., Grifoll, M. (2015). Bacterial PAH degradation in marine and terrestrial habitats. *Current Opinion in Biotechnology*, *33*, 95–102.

Bioleaching, biofilms, and biocorrosion 294/295

http://en.wikipedia.org/wiki/Bioleaching

http://en.wikipedia.org/wiki/Microbial_corrosion

http://en.wikipedia.org/wiki/Biofilm

Flemming, H. C. (2002) Biofouling in water systems – cases, causes and countermeasures, *Appl Microbiol Biotechnol.* 59, 629

Brandl, H. (2001) *Microbial Leaching of Metals*, in Biotechnology 2nd Edition, 191. Vol. 10, Rehm, H. (ed.). Wiley-VCH, ISBN 3–527–28320-X

Gadd, G. (2001) *Accumulation and Transformation of Metals by Microorganisms*, in Biotechnology 2nd Edition, 225. Vol. 10, Rehm, H. (ed.). Wiley-VCH, ISBN 3–527–28320-X

Sand, W. (2001) *Microbial Corrosion and its Inhibition*, in Biotechnology 2nd Edition, 265. Vol. 10, Rehm, H. (ed.). Wiley-VCH, ISBN 3–527–28320-X

Espinosa-Reyes, G., González-Mille, D. J., Iizaliturri-Hernández, C. A., Mejía-Saavedra, J., Cilia-López, V. G., Costilla-Salazar, R., Díaz-Barriga, F. (2014). Effect of mining activities in biotic communities of Villa de la Paz, San Luis Potosi, Mexico. *BioMed Research International*, *2014*, 165046.

The human genome 296/297

http://en.wikipedia.org/wiki/Human_genome

http://en.wikipedia.org/wiki/Human_Genome_Project

Functional analysis of the human genome 298/299

http://en.wikipedia.org/wiki/Fondation_Jean_Dausset-CEPH

http://en.wikipedia.org/wiki/Human_Genome_Diversity_Project

http://en.wikipedia.org/wiki/Single-nucleotide_polymorphism

http://en.wikipedia.org/wiki/Hapmap

http://en.wikipedia.org/wiki/1000_Genomes_Project

Green, R: (2001) *Genomics and Human Disease*, in Biotechnology 2nd Edition, 105. Vol. 5b, Sensen, C. (ed.). Wiley-VCH, ISBN 3-527-28328-5

Dellaire, G: (2001) *Genetic Disease*, in Biotechnology 2nd Edition, 61. Vol. 5b, Sensen, C. (ed.). Wiley-VCH, ISBN 3-527-28328-5

Mange, E. P., Mange, A. J. (2008) *Basic Human Genetics*, Rastogi Publications, ISBN 9788171333219

Overmann, K., Coolidge, F. (2013). Human species and mating systems: Neandertal-Homo sapiens reproductive isolation and the archaeological and fossil records. *Journal of Anthropological Sciences = Rivista Di Antropologia: JASS/Istituto Italiano Di Antropologia*, *91*, 91–110.

Individualized medicine 300/301

http://en.wikipedia.org/wiki/Personalized_medicine

http://en.wikipedia.org/wiki/ADME

Zanger, U. M., Klein, K. (2013). Pharmacogenetics of cytochrome P450 2B6 (CYP2B6): advances on polymorphisms, mechanisms, and clinical relevance. *Frontiers in Genetics*, *4*, 24.

Claus, S. P., Swann, J. R. (2013). Nutrimetabonomics:applications for nutritional sciences, with specific reference to gut microbial interactions. *Annual Review of Food Science and Technology*, *4*, 381–99.

Jørgensen, J. T. (2015). Companion diagnostics: the key to personalized medicine. *Expert Review of Molecular Diagnostics*, *15*(2), 153–6.

DNA analysis 302/303

http://en.wikipedia.org/wiki/DNA_profiling

http://en.wikipedia.org/wiki/Prenatal_diagnosis
http://en.wikipedia.org/wiki/Genetic_disorder#Single_gene_disorder

Gene therapy 304/305

http://en.wikipedia.org/wiki/Gene_therapy
Mountain, A. (2000), Gene therapy: the first decade *Trends Biotechnol* 18, 119
Wu, N., Ataai, M.M. (2000), Production of viral vectors for gene therapy applications *Curr Opin Biotechnol* 11, 205
Mountain, A: (1999) *Overview of Gene Therapy*, in Biotechnology Second, Completely Revised Edition, 383. Vol. 5a, Ney, U., Schomburg, D: (ed.). VCH Verlagsgesellschaft, ISBN 3-527-28315-3
Carter, B. (1999) *Viral Vectors for Gene Therapy* in Biotechnology Second, Completely Revised Edition, 395. Vol. 5a, Ney, U., Schomburg, D. (ed.). VCH Verlagsgesellschaft, ISBN 3-527-28315-3
Weir, N. (1999) *Non-Viral Vectors for Gene Therapy* in Biotechnology Second, Completely Revised Edition, 427. Vol. 5a, Ney, U., Schomburg, D. (ed.). VCH Verlagsgesellschaft, ISBN 3-527-28315-3
Edelstein, M.L., Abedi, M.R., Wixon, J. (2007). Gene therapy clinical trials worldwide to 2007–an update. *The Journal of Gene Medicine*, *9*(10), 833–42.
Overview: http://www.genetherapynet.com/types-of-gene-therapy.html

iPS cell therapy 306/307

http://en.wikipedia.org/wiki/Induced_pluripotent_stem_cell
Maeda, T., Lee, M.J., Palczewska, G., Marsili, S., Tesar, P.J., Palczewski, K., ... Maeda, A. (2013). Retinal pigmented epithelial cells obtained from human induced pluripotent stem cells possess functional visual cycle enzymes in vitro and in vivo. *The Journal of Biological Chemistry*, *288*(48), 34484–93.
Nakamura, Y., Matsuo, J., Miyamoto, N., Ojima, A., Ando, K., Kanda, Y., ... Sekino, Y. (2014). Assessment of testing methods for drug-induced repolarization delay and arrhythmias in an iPS cell-derived cardiomyocyte sheet: multi-site validation study. *Journal of Pharmacological Sciences*, *124*(4), 494–501.
Inoue, H., Nagata, N., Kurokawa, H., Yamanaka, S. (2014). iPS cells: a game changer for future medicine. *The EMBO Journal*, *33*(5), 409–17.
Komae, H., Sekine, H., Dobashi, I., Matsuura, K., Ono, M., Okano, T., Shimizu, T. (2015). Three-dimensional functional human myocardial tissues fabricated from induced pluripotent stem cells. *Journal of Tissue Engineering and Regenerative Medicine*.

Tissue Engineering 308/309

http://en.wikipedia.org/wiki/Tissue_engineering
Alépée, N. (2014). State-of-the-art of 3D cultures (organs-on-a-chip) in safety testing and pathophysiology. *ALTEX*, *31*(4), 441–77.
Baker, N., Boyette, L.B., Tuan, R.S. (2014). Characterization of bone marrow-derived mesenchymal stem cells in aging. *Bone*, *70*, 37–47.
Lynch, T.S., Patel, R.M., Benedick, A., Amin, N.H., Jones, M.H., Miniaci, A. (2015). Systematic Review of Autogenous Osteochondral Transplant Outcomes. *Arthroscopy: The Journal of Arthroscopic & Related Surgery: Official Publication of the Arthroscopy Association of North America and the International Arthroscopy Association*.
Mcheik, J.N., Barrault, C., Levard, G., Morel, F., Bernard, F.-X., Lecron, J.-C. (2014). Epidermal healing in burns: autologous keratinocyte transplantation as a standard procedure: update and perspective. *Plastic and Reconstructive Surgery. Global Open*, *2*(9), e218.
Paulsen, S.J., Miller, J.S. (2015). Tissue vascularization through 3D printing: Will technology bring us flow? *Developmental Dynamics: An Official Publication of the American Association of Anatomists*.
Zabala, M.E., Favre, J., Andriacchi, T.P. (2015). Relationship Between Knee Mechanics and Time Since Injury in ACL-Deficient Knees Without Signs of Osteoarthritis. *The American Journal of Sports Medicine*.

Drug screening 310/311

http://en.wikipedia.org/wiki/Pharmaceutical_drug

Gregersen, J. (1995) *Biomedicinal Product Development*, in Biotechnology 2nd Edition, 213. Vol. 12, Brauer, D: (ed.). Wiley-VCH, ISBN 3-527-28322-6

Harvey, A. L., Edrada-Ebel, R., Quinn, R. J. (2015). The re-emergence of natural products for drug discovery in the genomics era. *Nature Reviews Drug Discovery*, *14*(2), 111–29.

Prével, C., Kurzawa, L., Van, T. N. N., Morris, M. C. (2014). Fluorescent biosensors for drug discovery new tools for old targets–screening for inhibitors of cyclin-dependent kinases. *European Journal of Medicinal Chemistry*, *88*, 74–88.

Schadt, E. E., Buchanan, S., Brennand, K. J., Merchant, K. M. (2014). Evolving toward a human-cell based and multiscale approach to drug discovery for CNS disorders. *Frontiers in Pharmacology*, *5*, 252.

Walsh, K. B. (2015). Targeting cardiac potassium channels for state-of-the-art drug discovery. *Expert Opinion on Drug Discovery*, *10*(2), 157–69.

High-throughput sequencing (HTS) 312/313

http://en.wikipedia.org/wiki/High-throughput_sequencing#High-throughput_sequencing

Harrington, C. T., Lin, E. I., Olson, M. T., Eshleman, J. R. (2013). Fundamentals of pyrosequencing. *Archives of Pathology & Laboratory Medicine, 137*(9), 1296–303.

Jünemann, S., Sedlazeck, F. J., Prior, K., Albersmeier, A., John, U., Kalinowski, J., ... Harmsen, D. (2013). Updating benchtop sequencing performance comparison. *Nature Biotechnology, 31*(4), 294–6.

Makley, L. N., Gestwicki, J. E. (2013). Expanding the number of "druggable" targets: non-enzymes and protein-protein interactions. *Chemical Biology & Drug Design, 81*(1), 22–32.

Mayo, B., Rachid, C. T. C. C., Alegría, A., Leite, A. M. O., Peixoto, R. S., Delgado, S. (2014). Impact of next generation sequencing techniques in food microbiology. *Current Genomics, 15*(4), 293–309.

Nilakanta, H., Drews, K. L., Firrell, S., Foulkes, M. A., Jablonski, K. A. (2014). A review of software for analyzing molecular sequences. *BMC Research Notes, 7*(1), 830.

Paszkiewicz, K. H., Farbos, A., O'Neill, P., Moore, K. (2014). Quality control on the frontier. *Frontiers in Genetics*, 5, 157.

Proteomics 314/315

http://en.wikipedia.org/wiki/Proteomics

Dovichi, N., Hu, S., Michels, D., Zhang, Z., Krylov, S: (2001) *Proteome Analysis by Capillary Electrophoresis*, in Biotechnology 2nd Edition, 269. Vol. 5b, Sensen, C. (ed.). Wiley-VCH, ISBN 3-527-28328-5

Lottspeich, F. (2009). Introduction to proteomics. *Methods in Molecular Biology (Clifton, N. J.)*, *564*, 3–10.

Gold, L., Walker, J. J., Wilcox, S. K., Williams, S. (2012). Advances in human proteomics at high scale with the SOMAscan proteomics platform. *New Biotechnology*, *29*(5), 543–9.

Boersema, P. J., Kahraman, A., Picotti, P. (2015). Proteomics beyond large-scale protein expression analysis. *Current Opinion in Biotechnology*, *34C*, 162–170.

Transcriptomics (arrays) 316/317

http://en.wikipedia.org/wiki/Transcriptom ics

http://en.wikipedia.org/wiki/DNA_microarray

http://en.wikipedia.org/wiki/Protein_array

Tessier, D., Thomas, D., Brousseau, R. (2001) *A DNA Microarrays Fabrication Strategy*, in Biotechnology 2nd Edition, 227. Vol. 5b, Sensen, C. (ed.). Wiley-VCH, ISBN 3-527-28328-5

Bier, F. F., von Nickisch-Rosenegk, M., Ehrentreich-Förster, E., Reiss, E., Henkel, J., Strehlow, R., Andresen, D. (2008). DNA microarrays. *Advances in Biochemical Engineering/biotechnology*, *109*, 433–53.

Liu, D., Hoynes-O'Connor, A., Zhang, F. (2013). Bridging the gap between systems biology and synthetic biology. *Frontiers in Microbiology*, *4*, 211.

Valdes, A. M., Glass, D., Spector, T. D. (2013). Omics technologies and the study of human ageing. *Nature Reviews. Genetics*, *14*(9), 601–7.

Balmer, N. V, Leist, M. (2014). Epigenetics and transcriptomics to detect adverse drug effects in model systems of human development.

Basic & Clinical Pharmacology & Toxicology, *115*(1), 59–68.

Metabolics and metabolic engineering 318/319

http://en.wikipedia.org/wiki/Protein_array
http://en.wikipedia.org/wiki/Metabolic_engineering
Sahm, H. (1993) *Metabolic Design*, in Biotechnology Second, Completely Revised Edition, 189. Vol. 1, Sahm, H. (ed.). VCH Verlagsgesellschaft, ISBN 3-527-28337-4
Stephanopoulos, G. N., Aristidou, A. A., Nielsen, J. (1998) *Metabolic Engineering – Principles and Methodologies*, Academic Press, ISBN 0-12-666260-6.
Stephanopoulos, G. (2012). Synthetic biology and metabolic engineering. *ACS Synthetic Biology*, *1*(11), 514–25.
Gibbons, H., O'Gorman, A., Brennan, L. (2015). Metabolomics as a tool in nutritional research. *Current Opinion in Lipidology*, *26*(1), 30–4.
Wen, B., Zhu, M. (2015). Applications of mass spectrometry in drug metabolism: 50 years of progress. *Drug Metabolism Reviews*, 1–17.
Simeonidis, E., Price, N. D. (2015). Genome-scale modeling for metabolic engineering. *Journal of Industrial Microbiology & Biotechnology*, *42*(3), 327–38.

Synthetic biology 320/321

http://en.wikipedia.org/wiki/Synthetic_biology
http://en.wikipedia.org/wiki/Biobricks
http://biobricks.org
Hale, V., Keasling, J. D., Renninger, N., Diagana, T. T. (2007). Microbially derived artemisinin: a biotechnology solution to the global problem of access to affordable antimalarial drugs. *The American Journal of Tropical Medicine and Hygiene*, *77*(6 Suppl), 198–202.
Esvelt, K. M., Wang, H. H. (2013). Genome-scale engineering for systems and synthetic biology. *Molecular Systems Biology*, *9*, 641.
Farhi, M., Kozin, M., Duchin, S., Vainstein, A. (2013). Metabolic engineering of plants for artemisinin synthesis. *Biotechnology & Genetic Engineering Reviews*, *29*, 135–48.
Paddon, C. J., Keasling, J. D. (2014). Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nature Reviews. Microbiology*, *12*(5), 355–67.

Systems biology 322/323

http://en.wikipedia.org/wiki/Systems_biology
Loew, L. M., Schaff, J. C. (2001), The Virtual Cell: a software environment for computational cell biology. *Trends Biotechnol* 19, 401.
Kitano, H. (2002), Systems biology: a brief overview. *Science* 295, 1662.
Székely, T., Burrage, K. (2014). Stochastic simulation in systems biology. *Computational and Structural Biotechnology Journal*, *12*(20–21), 14–25.
Ritchie, M. D., Holzinger, E. R., Li, R., Pendergrass, S. A., Kim, D. (2015). Methods of integrating data to uncover genotype–phenotype interactions. *Nature Reviews Genetics*, *16*(2), 85–97.
Simeonidis, E., Price, N. D. (2015). Genome-scale modeling for metabolic engineering. *Journal of Industrial Microbiology & Biotechnology*, *42*(3), 327–38.

Bioinformatics: sequence and structure 324/325

http://en.wikipedia.org/wiki/Bioinformatics
http://en.wikipedia.org/wiki/Protein_Data_Bank
Wishart, D. (2001) *Bioinformatics: Tools for Protein Technologies*, in Biotechnology 2nd Edition, 325. Vol. 5b, Sensen, C. (ed.). Wiley-VCH, ISBN 3-527-28328-5
Zdobnov, E., Lopez, R., Apweiler, R., Etzold, T. (2001) *Bioinformatics: Using the Molecular Biology Data*, in Biotechnology 2nd Edition, 281. Vol. 5b, Sensen, C: (ed.). Wiley-VCH, ISBN 3-527-28328-5
Cygler, M., Matte, A., Schrag, J. (2001) *Bioinformatics: Structure Information*, in Biotechnology 2nd Edition, 345. Vol. 5b, Sensen, C: (ed.). Wiley-VCH, ISBN 3-527-28328-5
Berman, H. M., Gabanyi, M. J., Groom, C. R., Johnson, J. E., Murshudov, G. N., Nicholls, R. A., ... Minor, W. (2015). Data to knowledge: how to get meaning from your result. *IUCrJ*, *2*(1), 45–58.

Dorn, M., e Silva, M. B., Buriol, L. S., Lamb, L. C. (2014). Three-dimensional protein structure prediction: Methods and computational strategies. *Computational Biology and Chemistry*, *53*, 251–276.

Sonnhammer, E. L. L., Gabaldón, T., Sousa da Silva, A. W., Martin, M., Robinson-Rechavi, M., Boeckmann, B., ... Dessimoz, C. (2014). Big data and other challenges in the quest for orthologs. *Bioinformatics (Oxford, England)*, *30*(21), 2993–8.

Zanghellini, A. (2014). de novo computational enzyme design. *Current Opinion in Biotechnology*, *29*, 132–8.

Bioinformatics: functional analysis 326/327

http://en.wikipedia.org/wiki/KEGG
http://en.wikipedia.org/wiki/BRENDA
http://en.wikipedia.org/wiki/BLAST

Carbon Sources 328/329

http://www3.weforum.org/docs/WEF_FutureIndustrialBiorefineries_Report_2010.pdf

Kircher, M. (2012). How to turn industrial biotechnology into reality. *New Biotechnology*, *29*(2), 243–7.

Kircher, M. (2015). Residuals originating from the agri- and silvicultural, industrial, transport and municipal sectors. *New Biotechnology*, *32*(1), 97–9.

Biorefineries 330/331

http://en.wikipedia.org/wiki/Biorefinery

Kamm, B., Kamm, M. (2004). Principles of biorefineries. *Applied Microbiology and Biotechnology*, *64*(2), 137–45.

Ragauskas, A. J., Williams, C. K., Davison, B. H., Britovsek, G., Cairney, J., Eckert, C. A., ... Tschaplinski, T. (2006). The path forward for biofuels and biomaterials. *Science (New York, N. Y.)*, *311*(5760), 484–9.

Kamm, B., Kamm, M. (2007). Biorefineries–multi product processes. *Advances in Biochemical Engineering/biotechnology*, *105*, 175–204.

Stöcker, M. (2008). Biofuels and biomass-to-liquid fuels in the biorefinery: catalytic conversion of lignocellulosic biomass using porous materials. *Angewandte Chemie (International Ed. in English)*, *47*(48), 9200–11.

Rögner, M. (2013). Metabolic engineering of cyanobacteria for the production of hydrogen from water. *Biochemical Society Transactions*, *41*(5), 1254–9.

Forman, G. S., Divita, V. B., Han, J., Cai, H., Elgowainy, A., Wang, M. (2014). U. S. refinery efficiency: impacts analysis and implications for fuel carbon policy implementation. *Environmental Science & Technology*, *48*(13), 7625–33.

Safety in genetic engineering 332/333

http://en.wikipedia.org/wiki/Regulation_of_genetic_engineering
http://www.bvl.bund.de/EN/06_Genetic_Eng ineering/genetic_engineering_node.html

Simon, R., Frommer, W. (1993) *Safety Aspects in Biotechnology*, in Biotechnology Second, Completely Revised Edition, 835. Vol. 2, Pühler, A: (ed.). VCH Verlagsgesellschaft, ISBN 3-527-28312-9

Brauer, D., Bröker, M., Kellermann, C., Winnacker, E. (1995) *Biosafety in rDNA Research and Production*, in Biotechnology 2nd Edition, 63. Vol. 12, Brauer, D. (ed.). Wiley-VCH, ISBN 3-527-28322-6

Medley, T., McCammon, S. (1995) *Strategic Regulations for Safe Development of Transgenic Plants*, in Biotechnology 2nd Edition, 197. Vol. 12, Brauer, D: (ed.). Wiley-VCH, ISBN 3-527-28322-6

Registration of bioproducts 334/335

http://en.wikipedia.org/wiki/Drug_development
http://en.wikipedia.org/wiki/Food_and_Drug_Administration
http://en.wikipedia.org/wiki/European_Medicines_Agency
vfA Report 2014, www.vfa-bio.de

Ethics and public acceptance 336/337

Sinemus, K. (2007). Talking with the public–challenging the public scare. *Biotechnology Journal*, *2*(9), 1047.

Oeschger, M. P., Silva, C. E. (2007). Genetically modified organisms in the United States: implementation, concerns, and public percep-

tion. *Advances in Biochemical Engineering/biotechnology*, *107*, 57–68.
Einsele, A. (2007). The gap between science and perception: the case of plant biotechnology in Europe. *Advances in Biochemical Engineering/biotechnology*, *107*, 1–11.

Patents in biotechnology 338/339

http://en.wikipedia.org/wiki/Patent
http://www.wipo.int/portal/en/index.html
Schrell, A., Bauser, H., Brunner, H. (2007). Biotechnology patenting policy in the European Union–as exemplified by the development in Germany. *Advances in Biochemical Engineering/biotechnology*, *107*, 13–39.
Webber, P. M. (2007). The patenting of novel biotech products. *IDrugs: The Investigational Drugs Journal*, *10*(7), 459–62.

Biotechnology in an international comparison 340/341

Schmid, R., Chung, B., Jones, A., Saono, S., Scriven, J., Tsai, J. (1995) *Biotechnology in the Asian-Pacific Region*, in Biotechnology 2nd Edition, 369. Vol. 12, Brauer, D. (ed.). Wiley-VCH, ISBN 3-527-28322-6
Beppu, T. (2000), Development of applied microbiology to modern biotechnology in Japan *Adv Biochem Eng Biotechnol* 69, 41
Schmid, A., Hollmann, F., Park, J., Buhler, B. (2002), The use of enzymes in the chemical industry in Europe. *Curr Opin Biotechnol* 13, 359.
Ernst Young (2013) *Beyond Borders – The Global Biotechnology Report 2013. Ernst & Young 2013* www.ey.com/uk

Index

Picture Credits

Material from the following publications inspired some of the graphs in this work:

p. 3: Renneberg, R. (1990). *Biohorizonte* Urania Publisher, ISBN 3-332-00320-8

pp. 7, 11, 13, 15, 17: Black, J. (1999). *Microbiology - principles and explorations* Prentice Hall, ISBN 0-13-920711-2

pp. 23, 121: Schlegel, H. G. (1992). *Allgemeine Mikrobiologie* Thieme Publisher, ISBN 3-13-444607-3

pp. 37, 319: Stephanopoulos, G. N. et al. (1998). *Metabolic Engineering - principles and Methodologies* Academic Press, ISBN 0-12-666260-6

pp. 39, 41, 51: Knippers, R. (2001). *Molekulare Genetik* Thieme Publishers, ISBN 3-1347-7008-3

p. 55, 61: Brown, T. (1996) *Gentechnologie für Einsteiger* Spektrum Akademischer Verlag, ISBN 3-8274-0059-7

pp. 71, 73, 297, 301: Brown, T. A. (1999). *Genomes*, John Wiley Sons, ISBN 1-85996-201-7 (S. 73: the genetic map of Drosophila is from Arthur Sturtevant) (S. 297: data from the following publications have been used: Oliver et al. (1992). Nature *357*, 38 San Miguel et al. (1996). Science *274*, 765 Blattner et al. (1997). Science *277*, 1453)

p. 81, 237: Janeway, C., Travers, P. (1997) *Immunobiology*, Current Biology Publisher, ISBN 0-443-05964-0

p. 103: Buchholz, K., Kasche, V. (1997). *Biokatalysatoren und Enzymtechnologie*, VCH Publishers, ISBN 3-527-2831-7

p. 119: Ottman, N., Smidt, H., de Vos, W. M., Belzer, C. (2012). The function of our microbiota: who is out there and what do they do? *Frontiers in Cellular and Infection Microbiology*, *2*, 104.

p. 123: Präve, P. et al. (1994). *Handbuch der Biotechnologie*, R. Oldenbourg Publisher, ISBN 3-486-26223-8

p. 157: Heslot, H., Biochemie (1998). 80, 19

p. 169: Faber, K. (2000). *Biotransformations in Organic Chemistry*, Springer Publisher, ISBN 3-540-66334-7

pp. 183, 187, 191, 195: Uhlig, H. (1991), *Enzyme arbeiten für uns*, Hanser Publisher, ISBN 3-446-15702-6

p. 204: the map on MRSA is from: ECDC/EARS-Net: Proportion of Methicillin Resistant Staphylococcus aureus (MRSA) Isolates in Participating Countries in 2012. In: Antimicrobial resistance interactive database (EARS-Net) für das Europäisches Zentrum für die Prävention und die Kontrolle von Krankheiten. European Centre for Disease Prevention and Control (ECDC), 2012

p 229: Tuddenheim, E. G. D. (1997). in *Molecular Biology in Medicine*, Cox, T. M., Sinclair, J. (ed.). Blackwell Science Publisher, ISBN 0-632-02785-1

p 231: Bode, W., Renatur, H. (1997). Current Opinions in Structural Biology 6, 865

p 241: Mange, E. J., Mange, A. P. (1992). *Basic Human Genetics* Sinauer Associates, ISBN 0-87893-495-2p. 263 mass spectrum: Essentials of Glycobiology, 2nd ed.. Mulloy, B., Hart, G. W., Stanley, P. (2009). Structural Analysis of Glycans. Cold Spring Harbor Laboratory Press., http://www.ncbi.nlm.nih.gov/books/NBK1898/

pp. 269, 273: Schenkel, J. (1995). *Transgene Tiere*, Spektrum Akademischer Verlag, ISBN 3-86025-269-0

pp. 303, 317: Primrose, S. B. (1966). *Genomanalyse*, Spektrum Akademischer Verlag, ISBN 3-8274-0116-X

p. 309: Thude, S., Fraunhofer Institute for Bioprocess and Membrane Technology, Stuttgart, personal communication

p. 315: Lottspeich (1999). Angewandte Chemie Int. Ed. 38, 2476

p. 327: KEGG-generated metabolism: http://www.genome.jp/kegg-bin/show_pathway?map00020

p. 329: the map on biomass trade is based on the World Forum Biorefinery Report 2010, http://www3.weforum.org/docs/WEF_FutureIndustrialBiorefineries_Report_2010.pdf

p. 331: Ökoeffizienz Indigo: D Schulte (2000). *Nachr. aus der Chemie* 626

Copyright permission was obtained for the following photographs:

p. 17: Black, J. (1999). *Microbiology - principles and explorations,* Prentice Hall, ISBN 0-13-920711-2

p 19: Scenedesmus algae: Qingdao Institute for BioEnergy and Bioprocess Technology, VR China, p

p. 19: Dunaliella algae: Cognis - BASF Australia

p. 19: Neochloris abundans: Sapphire Co., USA, Copyright requested

p. 99: Atkinson e, B.t al., Biochemical Engineering and Biotechnology Handbook (1992), p. 459, plate C

p. 67: Epigenetics of homozygous twins, Abb. von Zachary A Kaminsky, Science Watch, http://sciencewatch.com/dr/erf/2010/10-augerf/10augerfKami/, copyright requestedp. 87 Horizontal paddle mixer for SSF: Nagel, F. J., Tramper, J., Bakker, M. S., Rinzema, A. (2001). Temperature control in a continuously mixed bioreactor for solid-state fermentation. *Biotechnology and Bioengineering*, *72*(2), 219-30, Wiley publishers

p. 163: *Mortierella alpinea:* Kyoto University, http://www.kyoto-u.ac.jp/en/research/forefronts/archives/shimizu.html

p. 239: CSIRO, Clayton, Australia, cited from Science 281, p 511 (1998)

p. 241: T. Cox et al., Molecular Biology in Medicine, plate 25 (1997)

p. 271: Nature 300 (5893) (1982) title, Macmillan Magazines Ltd.

p. 283: AJ Büchting, Biologie in unserer Zeit 28, p 16 (1998)

p. 283: G Krczal, Transgene Pflanzen in der landwirtschaftlichen Produktion, Nachr Chem Tech Lab 45, p 867 (1997)

p. 285: Susan Colburn/Monsanto, cited from Science 282, p 2178 (1998)

p. 285: Sainsbury Tomato Puree http://www.ncbe.reading.ac.uk/ncbe/gmfood/tomato.html

p. 331: Cargill Biorefinery Blair: http://greenchemicalsblog.blogspot.de/2009_03_01_archive.html

Photographs from the following Wikipedia articles were used

p. 15: Yeasts: http://de.wikipedia.org/wiki/Hefen
http://img2.wikia.nocookie.net/__cb20130326194932/getyourscienceon/images/e/e5/Saccharomyces-cerevisiae.jpg

p. 19: Spirulina: http://de.wikipedia.org/wiki/Spirulina
Euglena: http://en.wikipedia.org/wiki/Euglena
Botryococcus: http://en.wikipedia.org/wiki/Botryococcus
Chlorella: http://de.wikipedia.org/wiki/Chlorella

p. 31: Catalytic triad: http://en.wikipedia.org/wiki/Catalytic_triad, modifiziert

p. 35: Liposome: http://en.wikipedia.org/wiki/Lipid_bilayer

p. 41: Elektron microscopie: R. Knippers (2006) *Molekulare Genetik* 9. Aufl., Thieme-Verlag

p. 43: Aptamer complex from protein data base: http://www.rcsb.org/pdb/home/home.do

p. 47: Restriction enzyme complex from protein data base: http://www.rcsb.org/pdb/home/home.do

p. 67: Agouti mice: http://en.wikipedia.org/wiki/Randy_Jirtle

p. 163: Cocoa plant http://en.wikipedia.org/wiki/Cocoa_bean

p. 163: Cocoa butter http://en.wikipedia.org/wiki/Cocoa_butter

The structure models on pp 29, 43, 47, 175, 211, 223, 225, 229, 231, 233, 235, 237, 239, 241 are from the Protein Databank, http://www.rcsb.org/pdb/home/home.do

In spite of thorough researches, it may be possible that copyright owners could not in all cases be duly identified. If pertinent copyright ownership can be proven to the publisher, a remuneration according to the customary rates of scientific publishing will be paid.