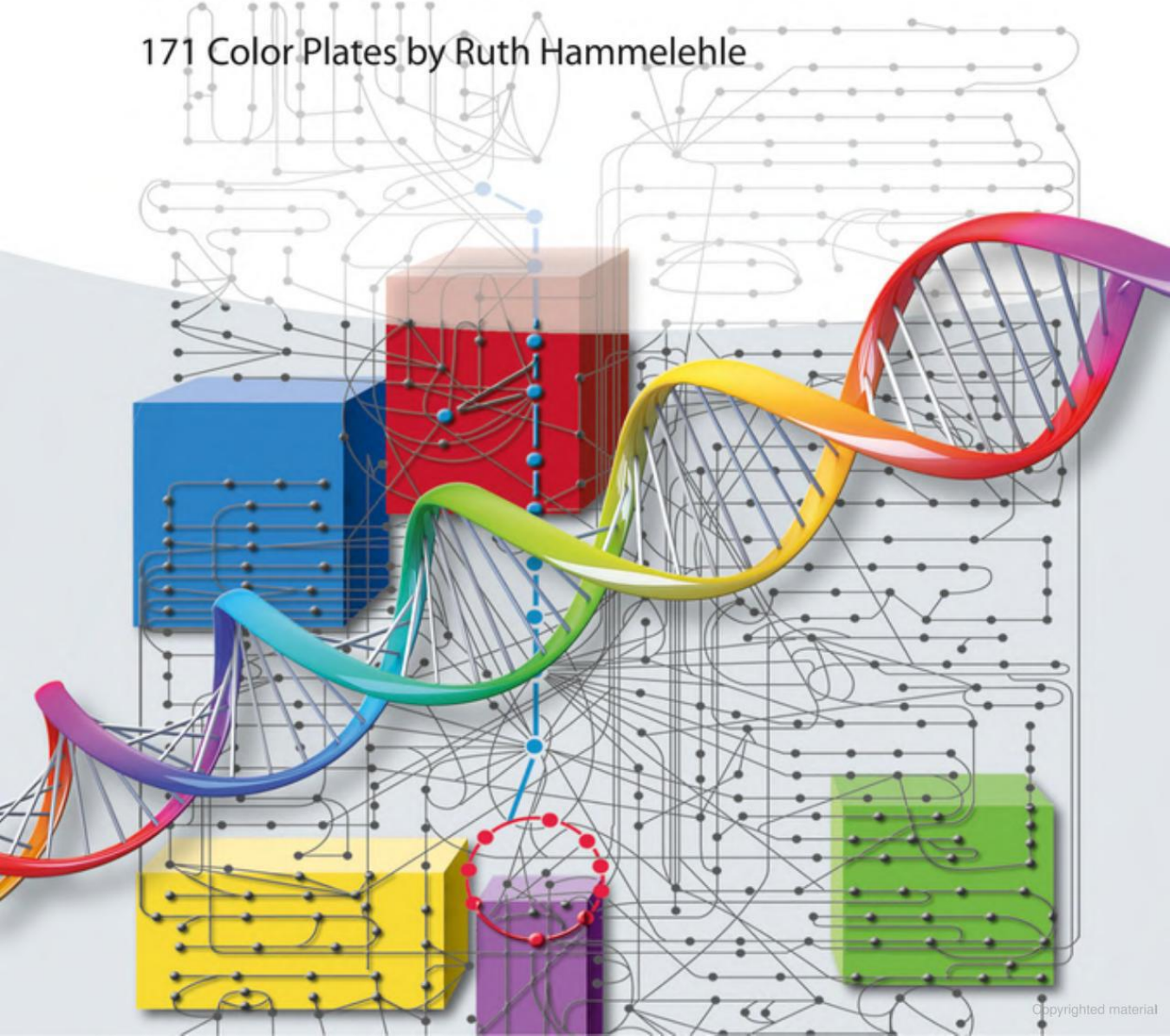
WILEY-VCH

Rolf D. Schmid and Claudia Schmidt-Dannert

Biotechnology

An Illustrated Primer

171 Color Plates by Ruth Hammelehle



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

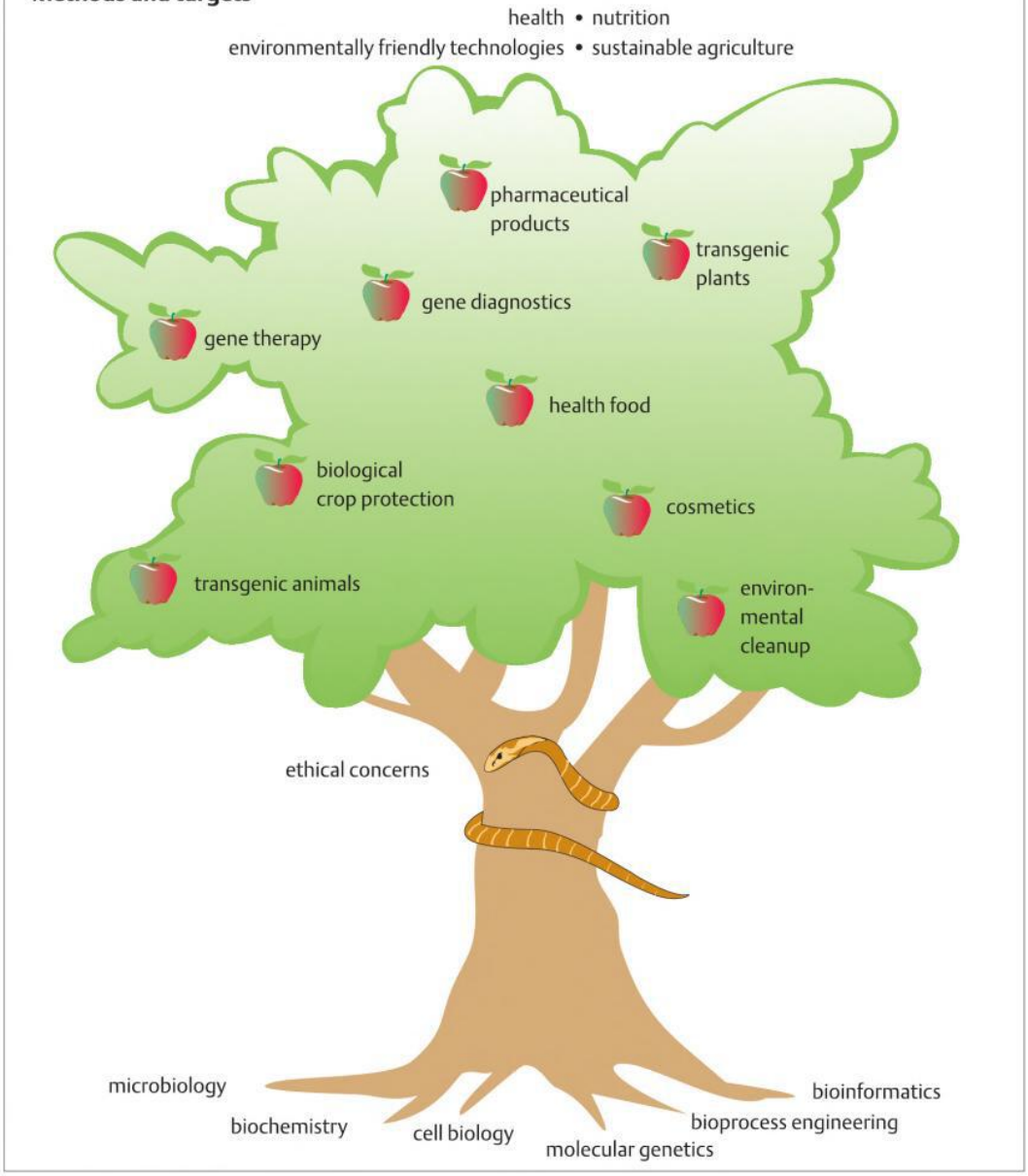
Biotechnology today

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.

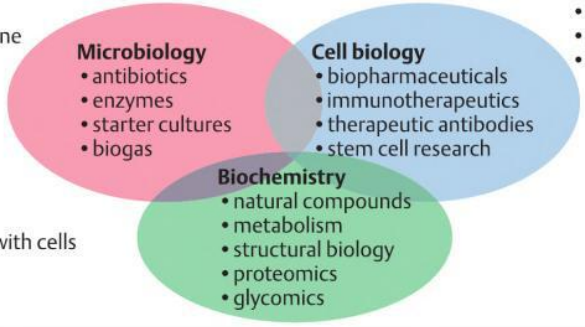
Methods and targets



Scientific foundations

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

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



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

Viruses

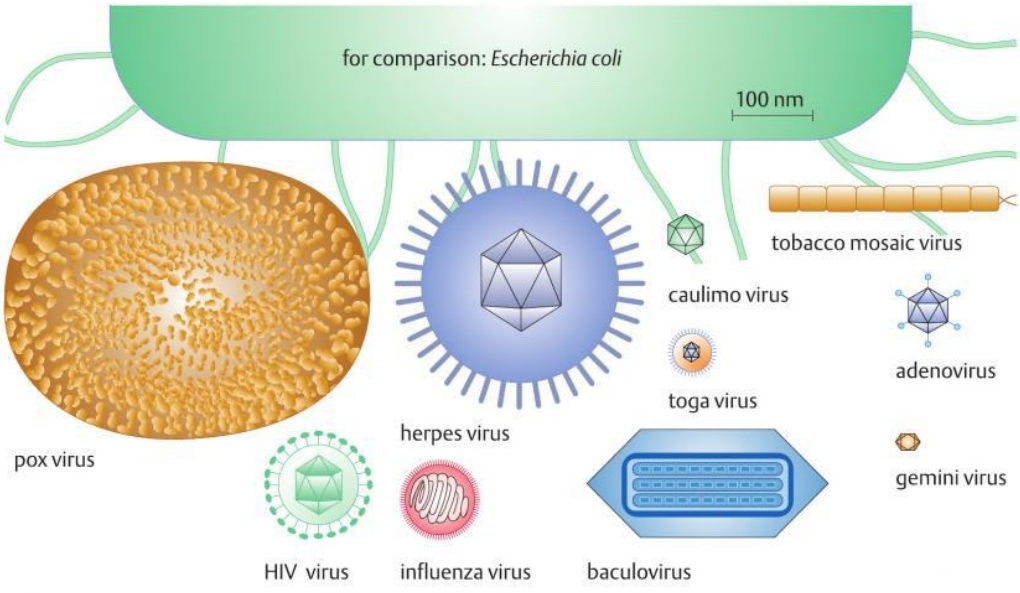
General. A virus is an infectious particle without indigenous metabolism. Its genetic program is written in either DNA or RNA, whose replication depends on the assistance of a living host cell. A virus propagates by causing its host to form a protein coat (capsid), which assembles with the viral nucleic 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.

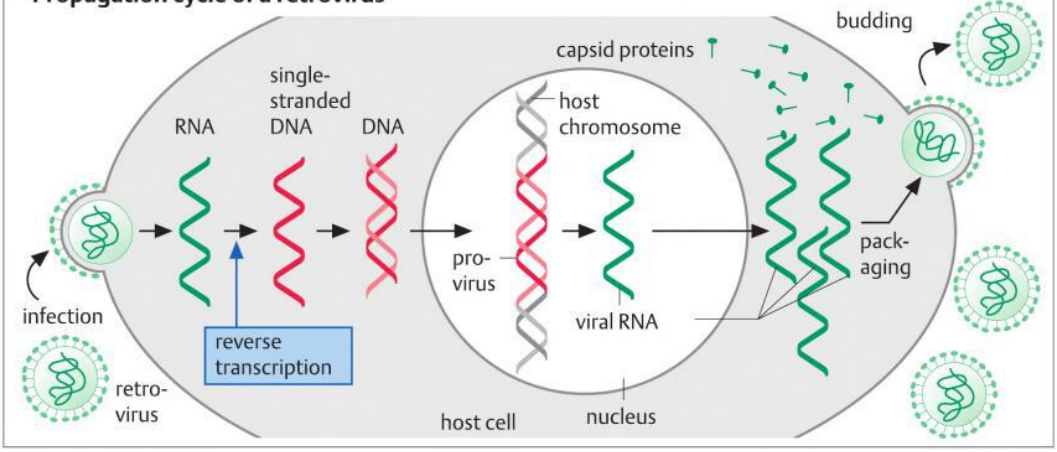
Forms



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.

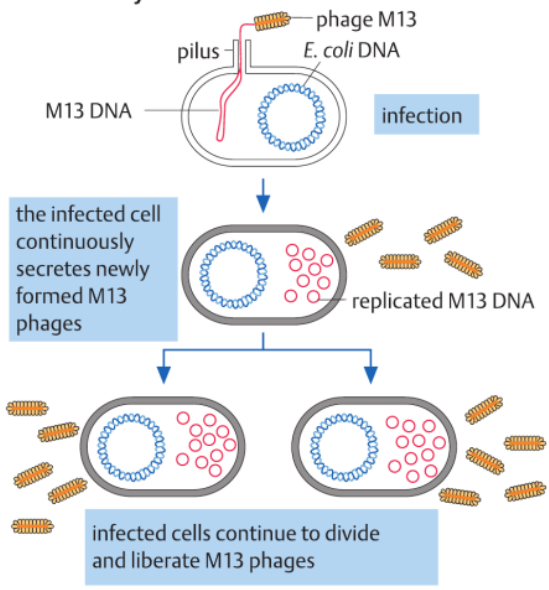
T Phages occur in 7 different types. For genetic engineering, two enzymes coded by T phage genomes are useful: the DNA ligase of T4, which links DNA fragments regardless of the quality of their ends (sticky or blunt), and the DNA polymerase of T7, which polymerizes DNA on a single strand DNA matrix; 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.

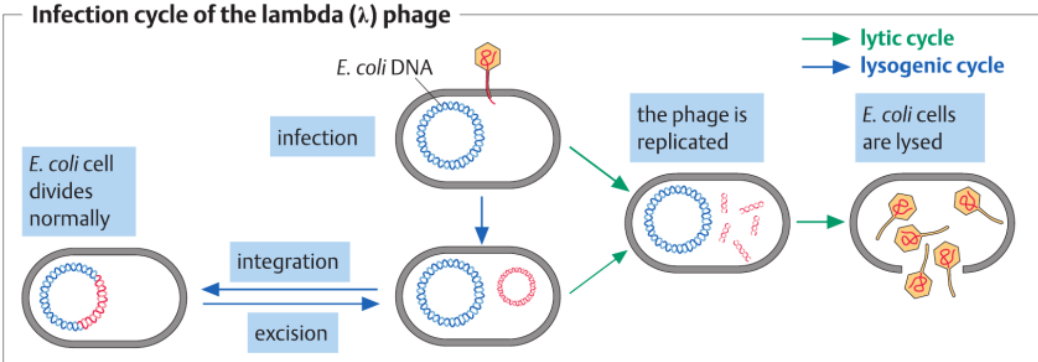
E. coli phages (select)

name	form	genetic material
T2 and T4	phage head, collar, tail (core, hull), basal sheet with spikes, tail fibers; 95,0 nm	DNA (double-stranded)
T7		DNA (double-stranded)
lambda (λ)		DNA (double-stranded)
M13	6 nm; 900 nm	DNA (single-stranded)

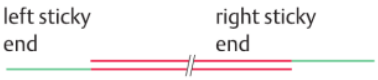
Infection cycle of M13



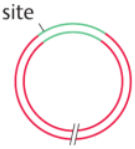
Infection cycle of the lambda (λ) phage

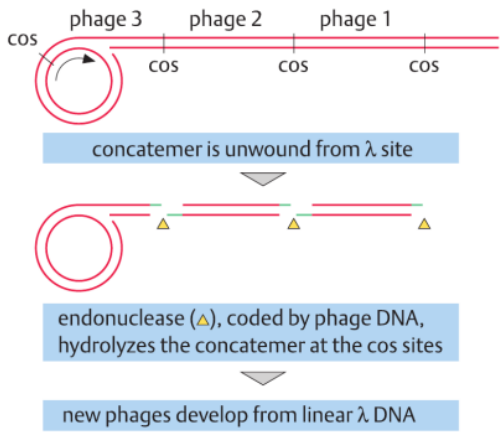


λ DNA in linear form

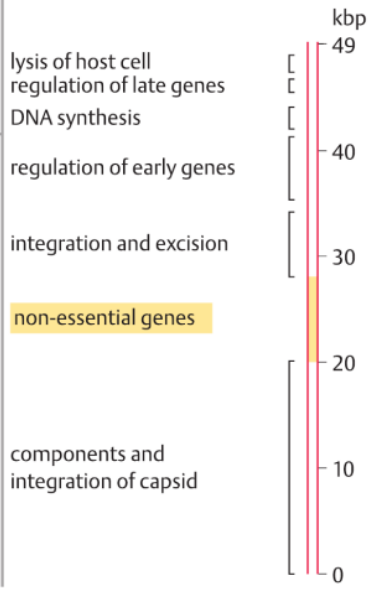


λ DNA in circular form





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).

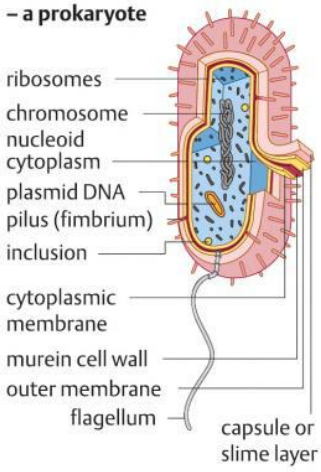
Eubacteria are unicellular organisms that propagate by cell division. Their cell diameter is usually on the order of 1 μm. They have no cell nucleus, and their chromosomal DNA is formed into a tangle, the nucleoid. Frequently, part of their genetic makeup occurs on nonchromosomal genetic elements, the 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).

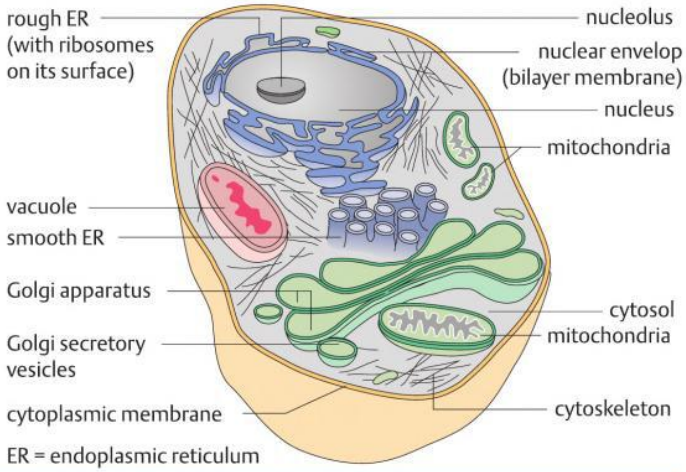
Yeasts and fungi are eukaryotic organisms and so far constitute the largest group of cultivatable microorganisms: about 70,000 different strains have been taxonomically classified. In contrast to prokaryotes, they contain a cell nucleus and other subcellular functional units, and their cell wall is made 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**

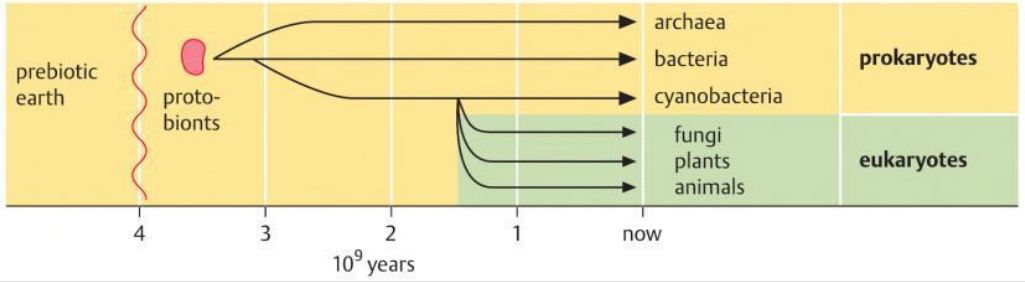


***Saccharomyces cerevisiae* – a eukaryote**



	E. coli	*S. cerevisiae*	for comparison: plant and animal cells
cell nucleus, organelles	no	yes	yes
diameter [µm]	~ 1	~ 10	~ 100
volume [µm³]	~ 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.

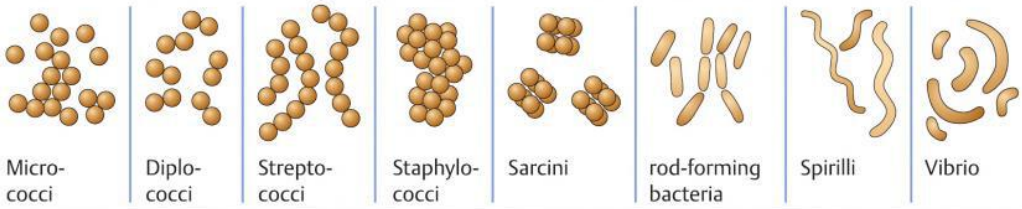
Phage typing: the susceptibility to phages can also be used for taxonomic identification,

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.

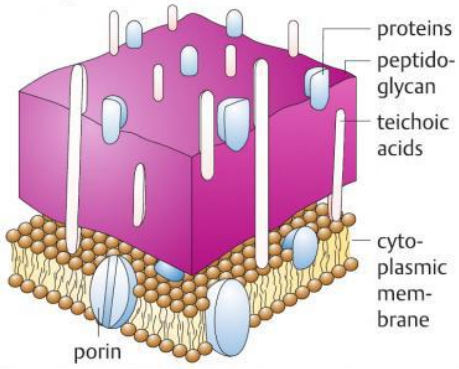
Genome sequencing. As of 2013, genome sequences for some 2,100 bacteria and over 140 archaea are completed. This includes many genomes of human pathogens such as *Mycobacterium tuberculosis*. The analysis of microbial genomes has shown that many variations of metabolic pathways exist, which can be exploited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

Forms of unicellular bacteria

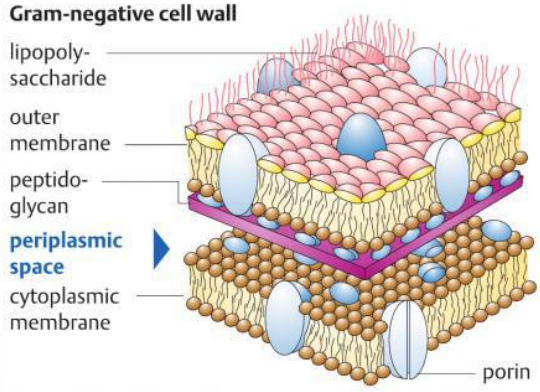


Cell wall composition and Gram-staining

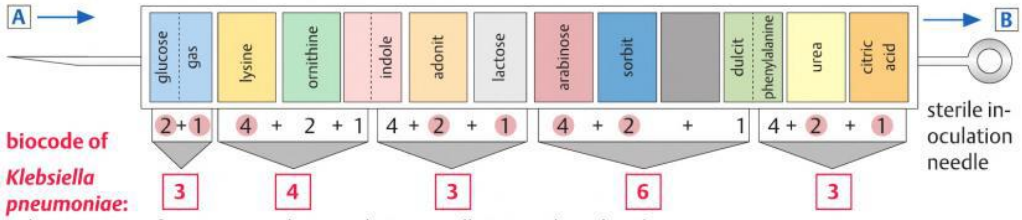
Gram-positive cell wall



Gram-negative cell wall

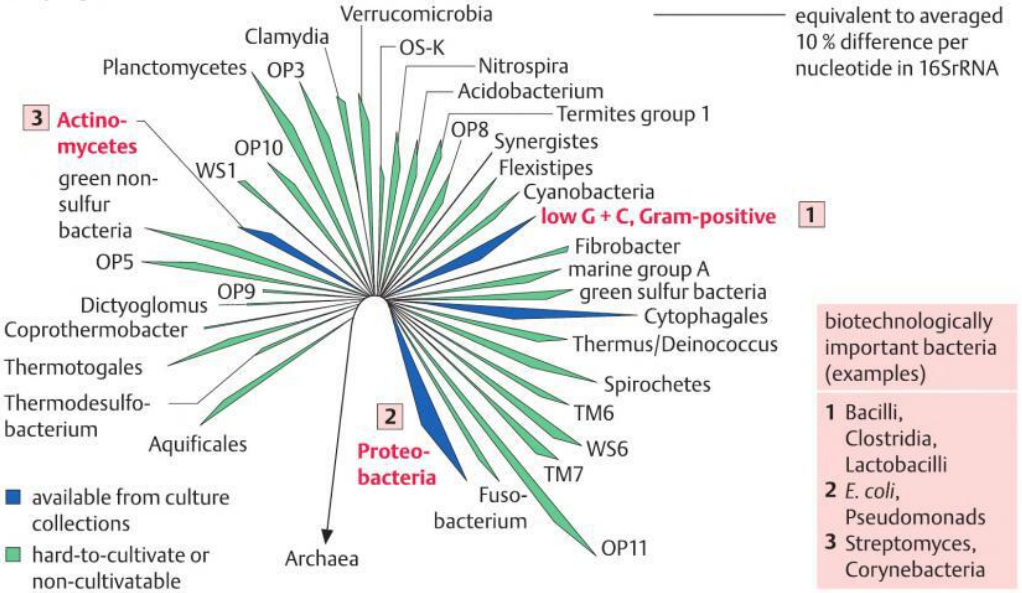


Biochemical characterization



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

Phylogenetics and cultivation



biotechnologically important bacteria (examples)

1 Bacilli, Clostridia, Lactobacilli
2 *E. coli*, Pseudomonads
3 Streptomyces, Corynebacteria

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*.

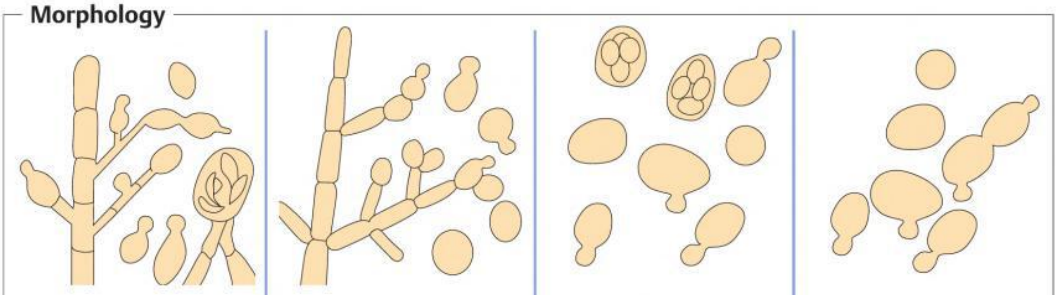
Saccharomyces cerevisiae (synonyms: baker's yeast, brewer's yeast, yeast) (→120) can propagate in either a haploid or diploid manner, thus providing an excellent organism for genetic investigations. Haploid laboratory strains belong to one of two mating types (*MATa* or *MATα*), which can only mate reciprocally. 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	Candida	Saccharomyces	Torulopsis
forms ascus	–	forms ascus	–
forms mycelium	forms mycelium	–	–
forms buds	forms buds	forms buds	forms buds

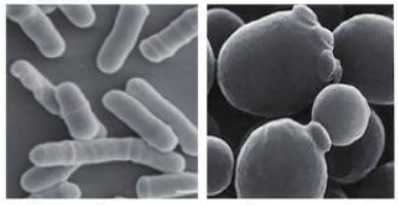
genetic

	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

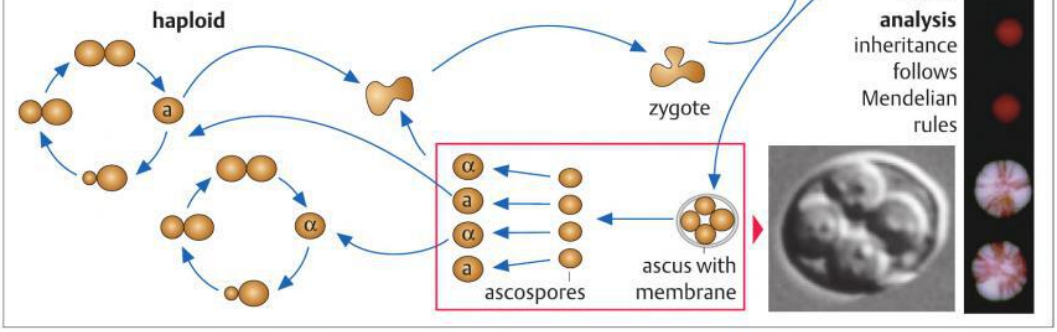
Yeasts



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.

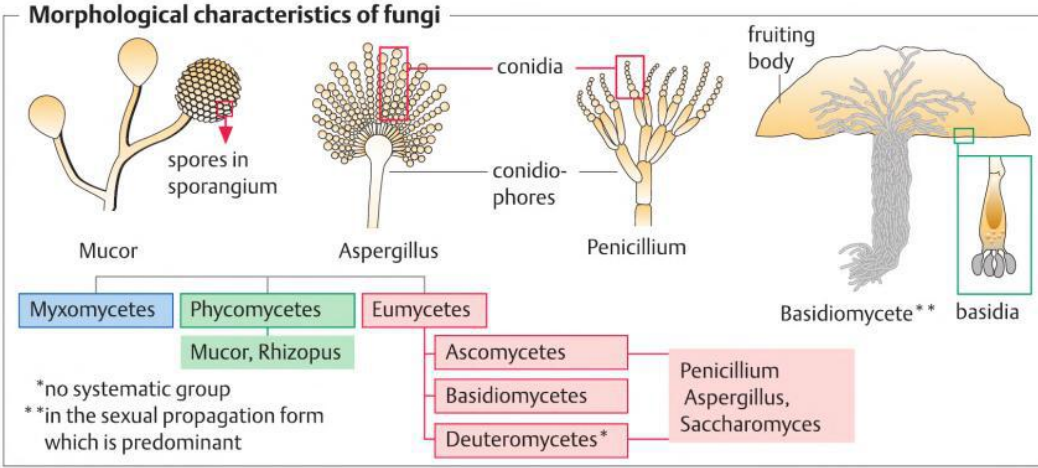
Reproduction forms. The reproduction of fungi follows highly diverse patterns, which are described here using the Ascomycetes as an example. The cell mass (thallus) consists of a mycelium that is made up of hyphae. During asexual reproduction, the conidiophores, which form at the top of the mycelium, 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.

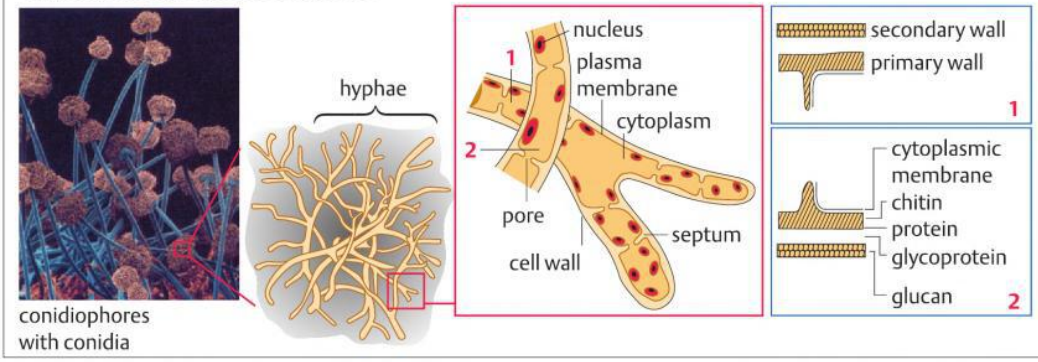
Aspergillus nidulans differs from *Penicillium* in the form of its conidia. Its genome contains 30.5 Mbp. *A. oryzae* is used for industrial production of extracellular enzymes (→172) and is a favorite host organism for producing recombinant enzymes from other eukaryotes. Various *Aspergillus* strains 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*.

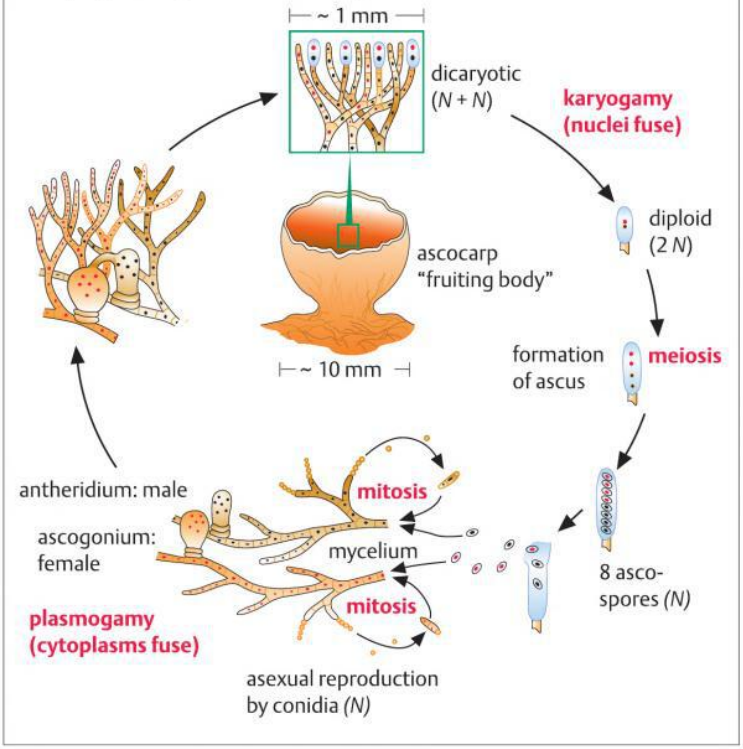
Morphological characteristics of fungi



Aspergillus niger, an ascomycete

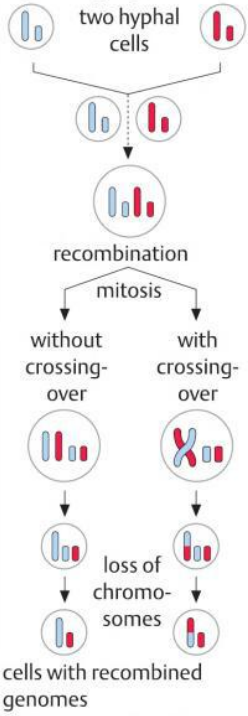


Propagation cycle of an ascomycete



Parasexual breeding

e.g., of Aspergillus



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*.

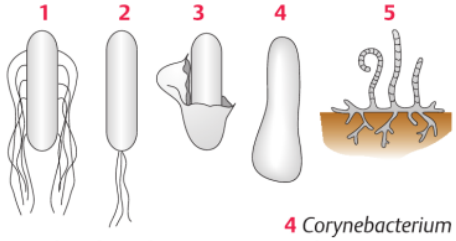
Escherichia coli is a saprophyte in the large intestine of mammals and belongs to the Enterobacteriaceae group. It forms rods that carry flagella. The cell wall stains Gram-negative: it encloses two membranes that include a periplasmic space. Under anaerobic growth conditions, *E. coli* generates energy 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

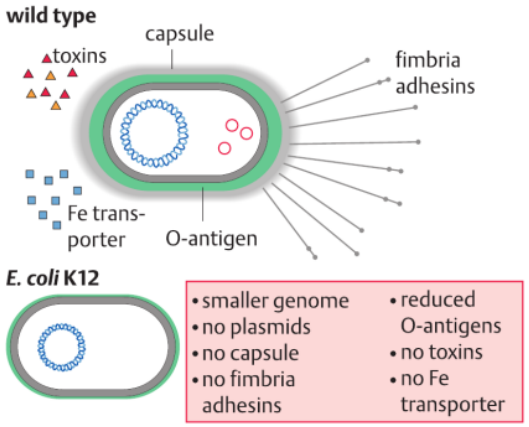


1 *Escherichia coli*
2 *Pseudomonas putida*
3 *Bacillus subtilis* (germinating from spore)
4 *Corynebacterium glutamicum*
5 *Streptomyces coelicolor* (with conidia)

	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

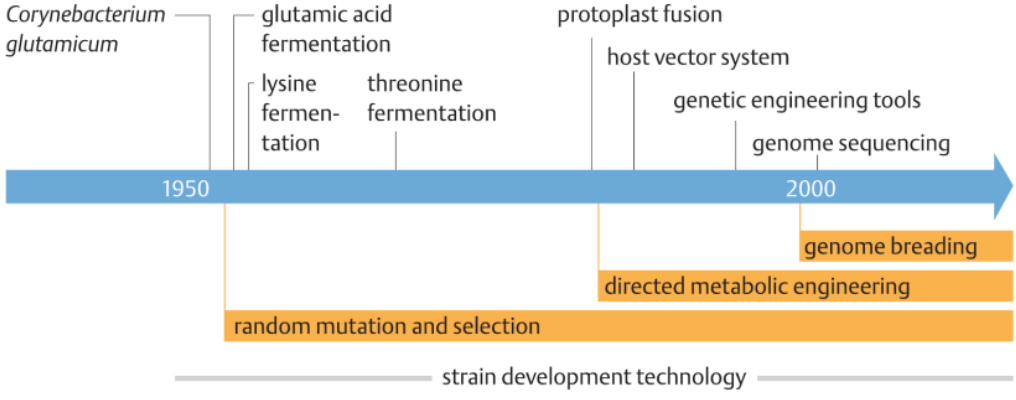
E. coli K12 modifications



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.

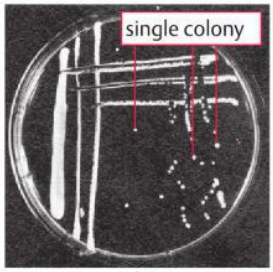
Pure cultures are obtained from culture collections or from their natural habitats (soil, water, food, other organisms) using enrichment cultures. The preferred method for obtaining a pure culture is the streak plate method, in which a mixed culture is spread over the surface of a sterile nutrient agar (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.

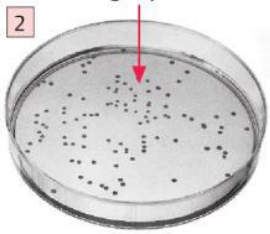
Pure cultures

1 streak plate method using nutrient agar



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

2

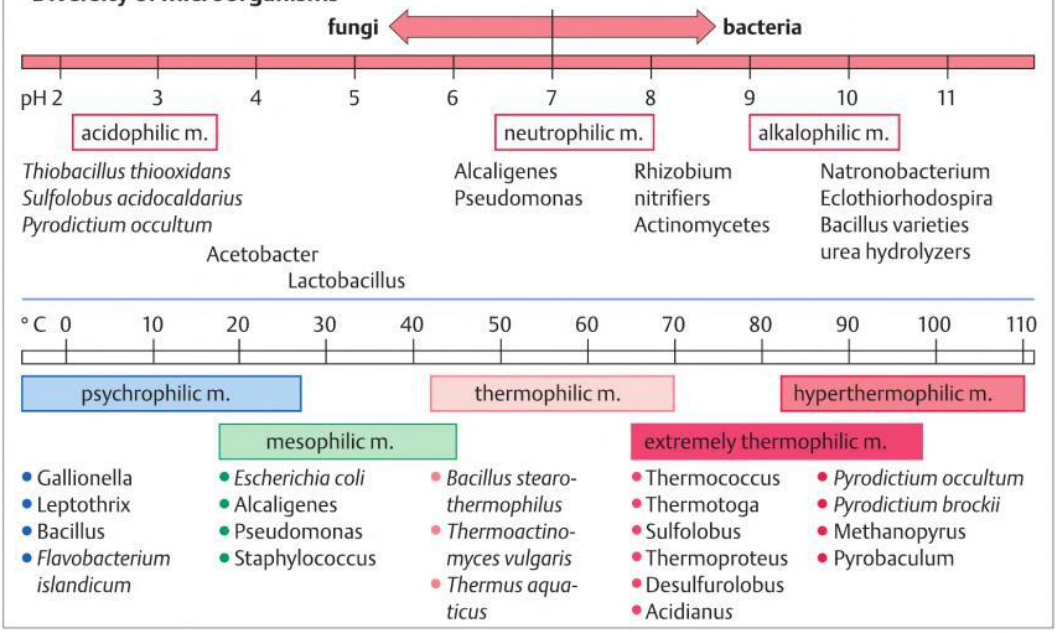


Enrichment cultures (examples)

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

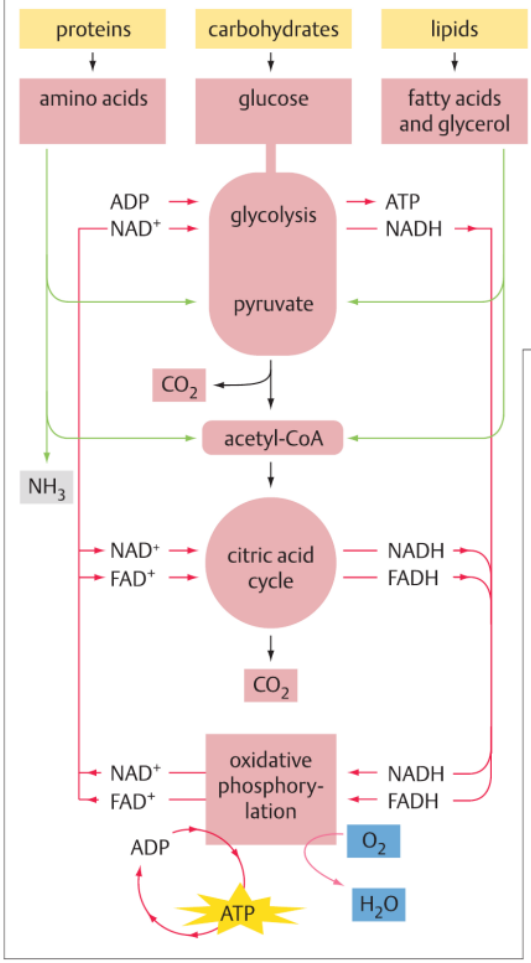
General. Microorganisms isolated from environmental samples rarely exhibit all the properties that are required in a technical application. Thus, they are usually optimized by a series of mutation and selection steps. The targets of strain improvement are usually: 1) to increase the yield of the desired 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).

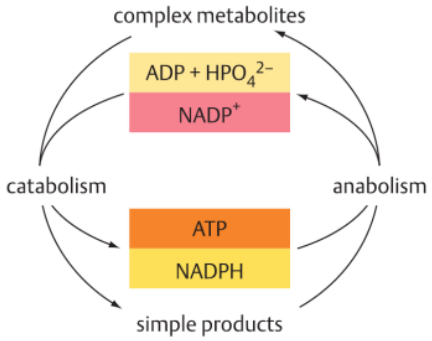
Mutation. The spontaneous mutation frequency (changes in DNA sequence due to natural mutation events and errors during replication) is on the order of 10^{-7} for a gene (1,000 bp) of normal stability. Most mutations remain silent or they revert genetically or functionally or by DNA repair mechanisms 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.

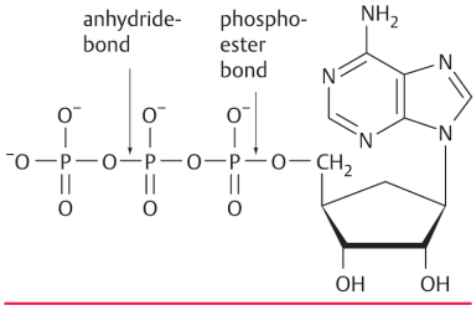
Selection in submersed culture. Continuous fermentation has also been used for selecting microorganisms. A pure culture of a microorganism is grown in a chemostat in the presence of a mutagenic agent and subjected to selective pressure, e. g., by gradually replacing a good carbon source (→88) with a 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.

Metabolites





High-energy compounds

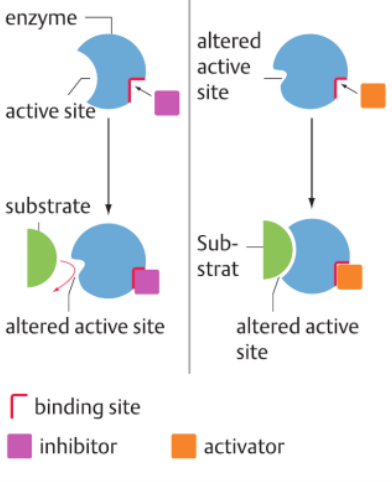


ATP ADP AMP adenosine

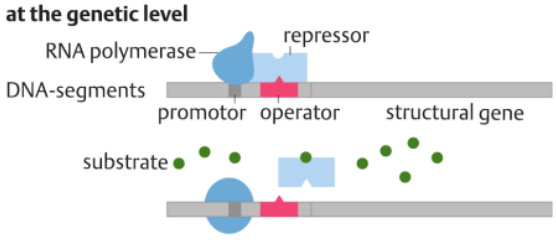
compound	free enthalpy $\Delta G^{0'}$
phosphenolpyruvate	−61,9 (kJ • mol^{-1})
ATP (→AMP + PP_i)	−45,6
ATP (→ADP + P_i)	−30,5
ADP	−19,2
glucose-6-phospate	−13,8

Regulation

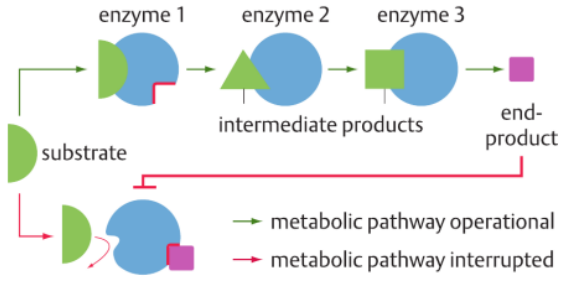
at the enzyme level



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.

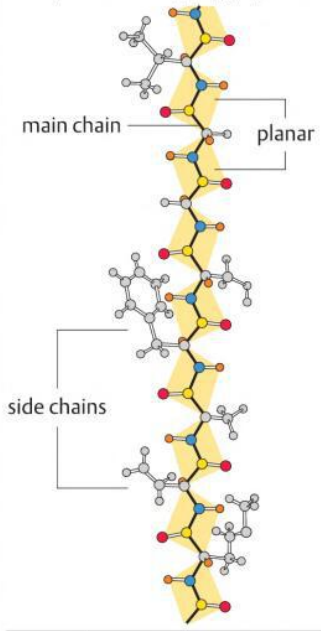
Proteins consist of one or several peptide chains that form a unique structure, mostly by non-covalent interactions. Globular proteins consist of ~31 % α-helix- and ~28 % ß-folds. The remaining ~40 % of the sequence form *random coil*s and turns. ß-turns are composed of 4 amino acids with glycine in 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.

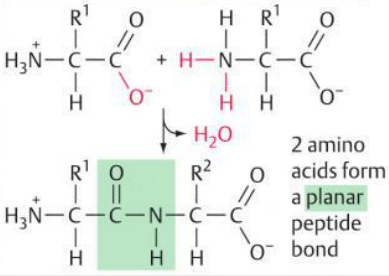
The protein-forming ("proteinogenic") amino acids

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

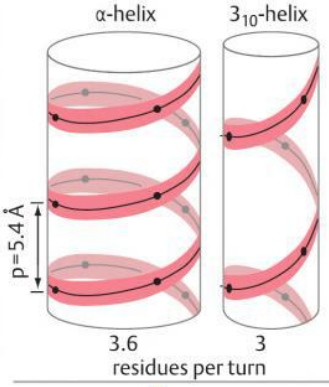
pi = isoelectric point

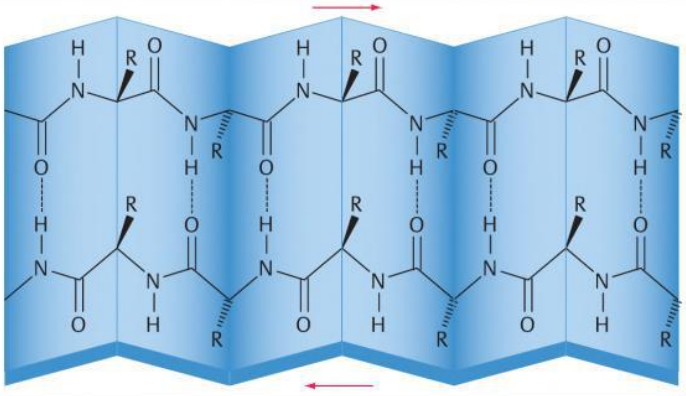
Peptide bond/polypeptides

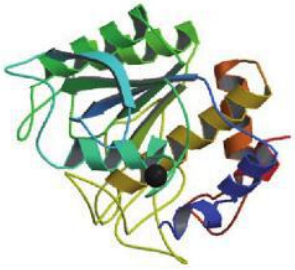




α-helix and β-pleated sheet







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

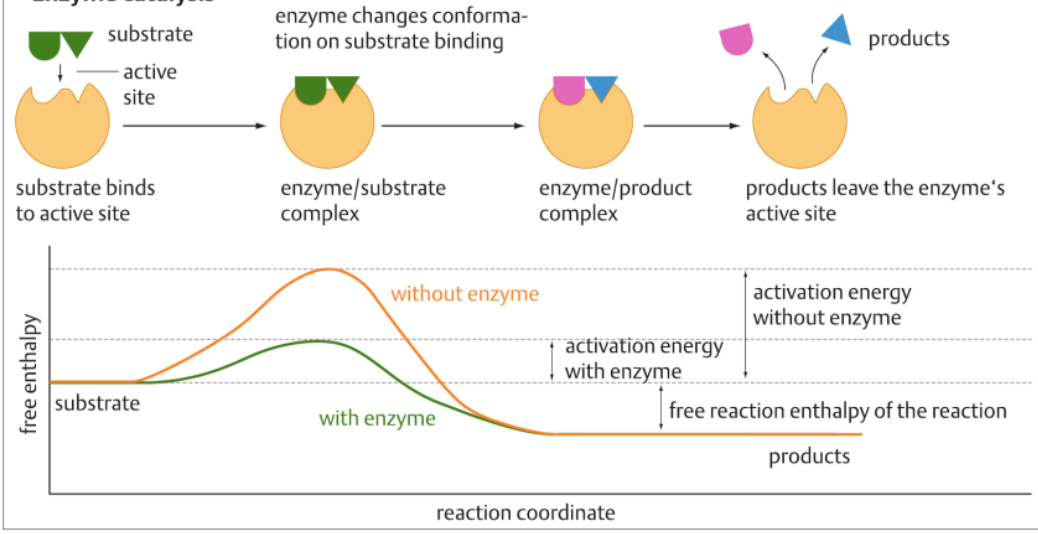
Enzyme (E) + Substrate (S) ⇋ Enzyme-Substrate or Michaelis-Complex (ES) ⇋ Product (P) + Enzyme (E), with the equilibrium assumption $dES/dt = 0$

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

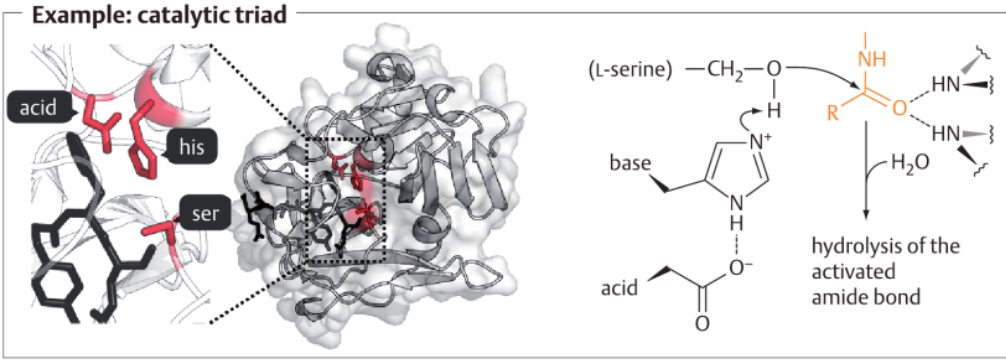
$$k_{cat} = V_{max}/E_T$$

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.

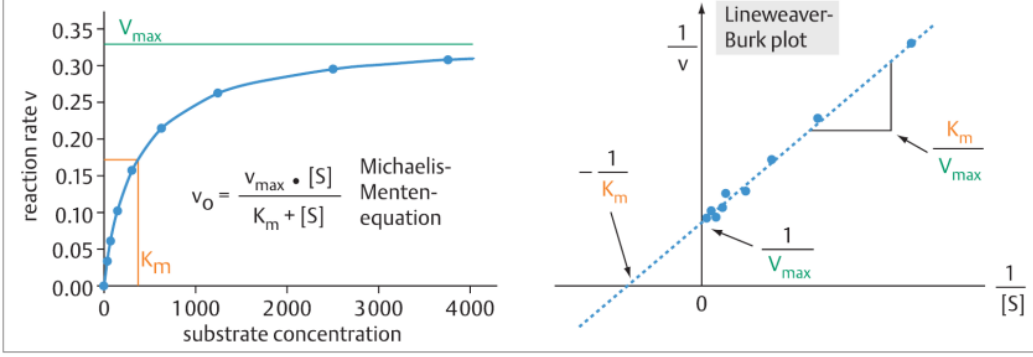
Enzyme catalysis



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

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.

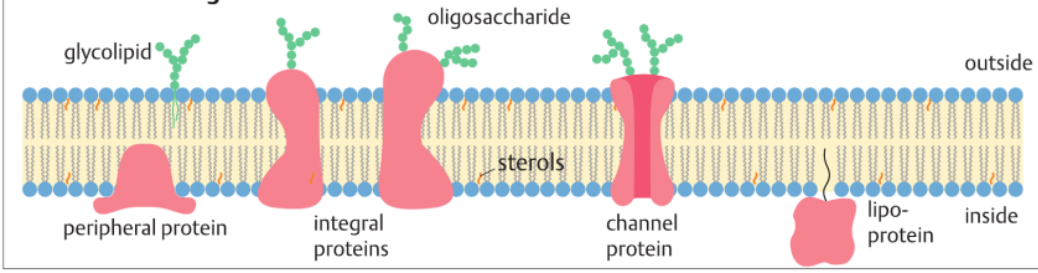
Biological membranes determine the "inside" and "outside" of cells, and also of most organelles inside eukaryotic cells (mitochondria, plastids, peroxisomes etc.). Their lipophilic components (e. g., phospholipids) diffuse very slowly between the two layers of the double membrane ("flip-flop", timescale: 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.

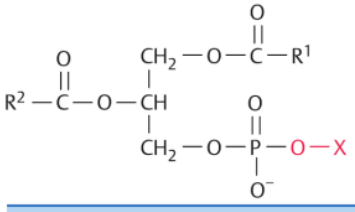
The most frequent natural lipids

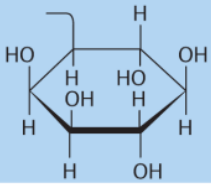
symbol	trivial name	common name	m.p. (°C)
saturated fatty acids			
12:0	lauric acid	dodecanoic acid	44.2
14:0	myristic acid	tetradecanoic acid	52
16:0	palmitic acid	hexadecanoic acid	63.1
18:0	stearic acid	octadecanoic acid	69.6
20:0	arachic acid	eicosanoic acid	75.4
unsaturated fatty acids (all double 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

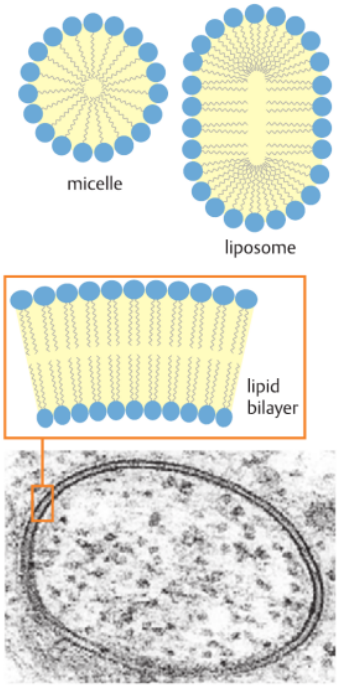


Glycerophospholipids



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



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.

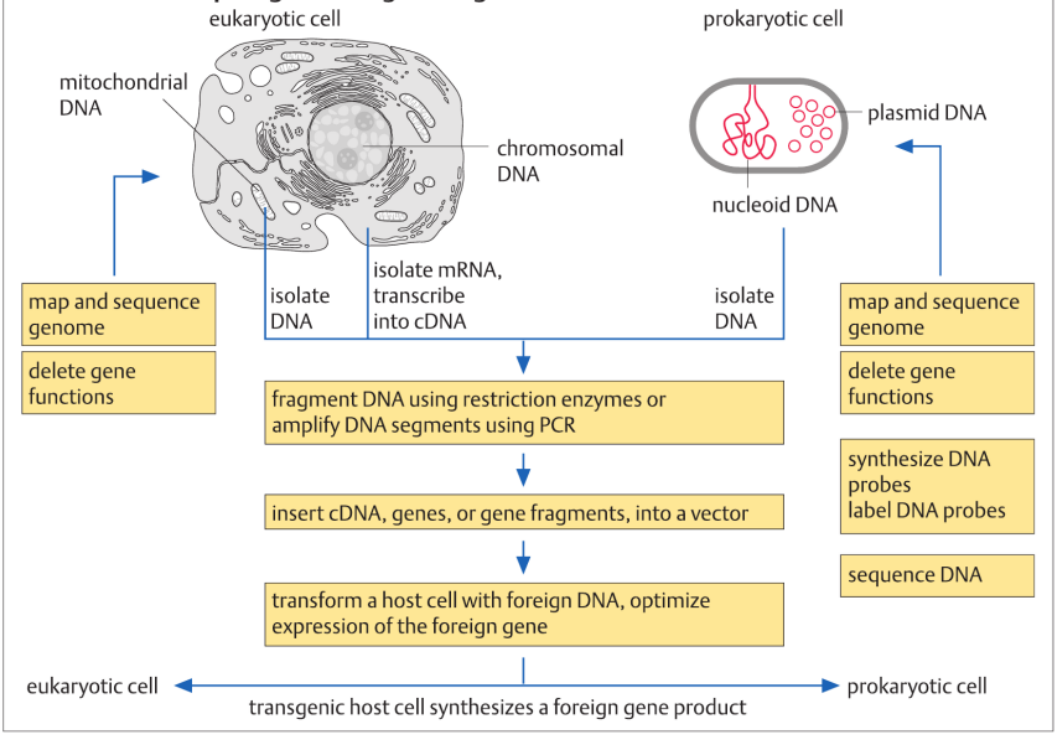
Aptamers are artificial nucleic acid ligands that bind with high affinity to hydrophilic molecules such as peptides and drugs. In the SELEX process (systematic evolution of ligands by exponential enrichment), vast combinatorial libraries of synthetic nucleic acids (10^{14}–10^{15} different molecules) 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.

Fundamental steps in genetic engineering



Cloning of eukaryotic genes

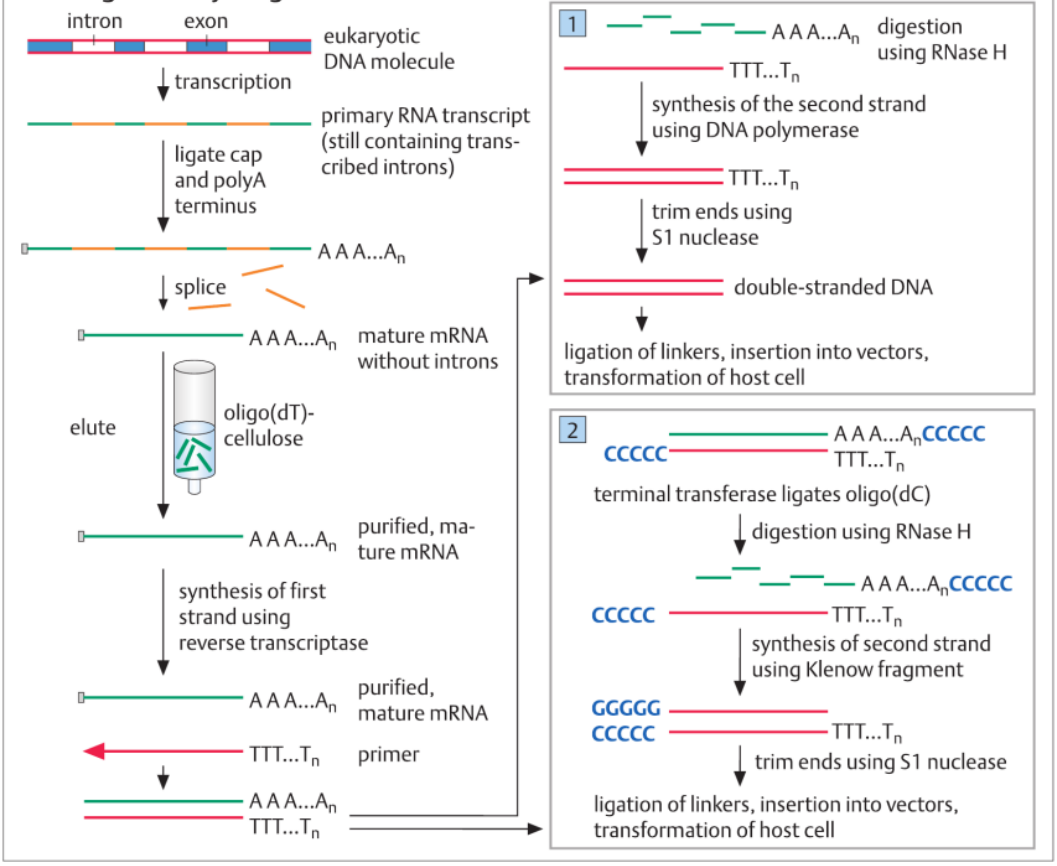



image
not
available


image
not
available