1
Biopharmaceutical Production: Value Creation, Product Types, and Biological Basics Introduction

1.1
Role of Production in Pharmaceutical Biotechnology

Over recent years, pharmaceutical biotechnology has developed very dynamically. An important driver for this success has been the enormous increase of scientific know-how in the areas of genetics and immunology, which has created huge expectations for the development of innovative medicinal treatments.

The scientific pioneer spirit has been fueled by public and private sponsorship, resulting in a biotechnological landscape that has long been dominated by highly innovative, venture capital-based, small- and mid-size companies. However, before patients can benefit from scientific achievements, it is necessary that the identified molecule is transformed into a medicine – fit for achieving the therapeutic target – and tested in comprehensive trials in the field. The production of such a medicine has to be carried out in officially licensed, often tailor-made technical manufacturing facilities.

From project to product

This path from project to product usually lasts several years and is associated with enormous costs and risks. On average, the development costs of a new compound are in the region of US$ 500–1000 million and only 10% of all projects that enter clinical trials find their way into the market.

Owing to these immense investments in drug development, the costs of drug manufacturing often seem acceptable, particularly as the costs are absorbed by sales of the marketed drug in the same accounting period; however, safe and efficient product supply is the cornerstone of a company's success. In biotechnology, the overlap between development and market launch is particularly intensive, motivating companies to take care of manufacturing early on:

- Many targets of process development result from requirements of large-scale manufacturing.
- The classical separation of development (pre-marketing) and production (post-marketing) does not work for biologics, as both the

Manufacturing of Pharmaceutical Proteins: From Technology to Economy, Third Edition. Stefan Behme.
© 2022 WILEY-VCH GmbH. Published 2022 by WILEY-VCH GmbH.

COPYRIGHTED MATERIAL

manufacturing process and plant are factors that determine the quality of the final medicinal product.

- Production is the basis for long-term market supply. Decisions about capital investment or outsourcing of manufacturing mostly have to be taken long before the market launch of the product.
- Biotechnological processes are much more difficult to control than small-molecule preparations. The limited ability to monitor and characterize the product results in increased manufacturing risks.

Significance of production in the value chain

The main target of production is to supply the product safely and cost-efficiently. It is positioned between the development and marketing of a product. Figure 1.1 illustrates its significance in the value chain.

The chain starts with research that has a clear focus on the identification of targets, which involves analyzing the interaction between the biochemical molecule and its potential therapeutic functionality. In the subsequent development phase, a process for the scale-up and more consistent manufacturing of the molecule is designed. Here, the target structure is developed into a pharmaceutical form and tested in animals and humans as to its safety and efficacy. Once this is achieved, production kicks in, taking care of a high-quality and profitable product supply, addressing the following main tasks:

- When, where, and in what quantities should the drug be produced? (*Production concept*)

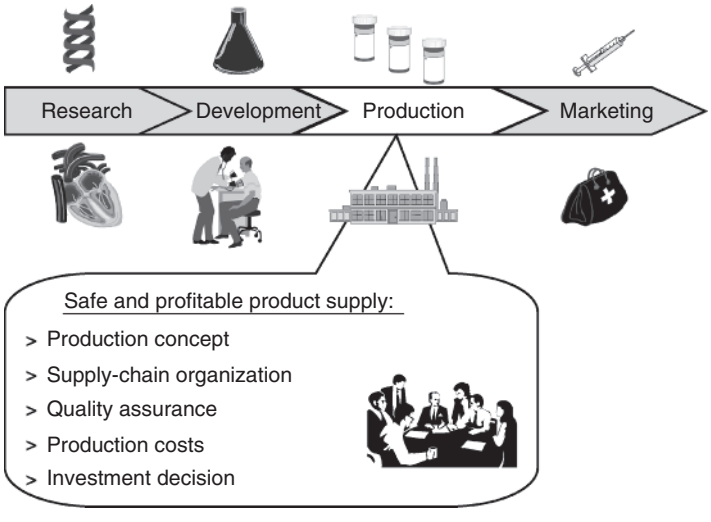

Figure 1.1 Role and tasks of production.

- How should market supply be organized? (*Supply-chain organization*)
- How should the quality of the product and Good Manufacturing Practice (GMP) compliance be assured? (*Quality assurance*)
- What are costs of manufacturing and how can these costs be controlled? (*Manufacturing costs*)
- How attractive is an investment in one's own facilities? (*Investment decision*)

The marketing of the product stands at the end of the value chain; from this position, essential goals are formulated for production: supply safety and cost efficiency.

Production is interdisciplinary

The integrated position of production in the value chain results in interdisciplinary tasks that are best treated by multilateral teams managed by experts in different disciplines such as biology, engineering, chemistry, economics, law, pharmacy, and medicine.

Figure 1.2 shows the subject areas that are important for the understanding and control of production processes and workflows. This volume provides an overview of these subject areas, while special emphasis is given to the interaction between these areas.

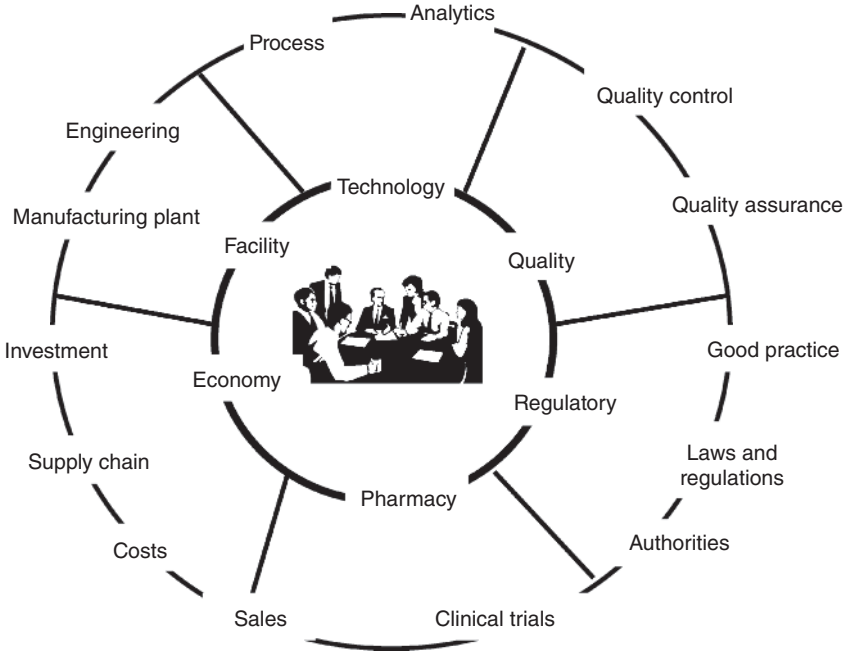

Figure 1.2 Subject areas in production. Inner circle = sections of this book; outer circle = subject areas treated in the sections.

Following this introductory part, Part II, "Technology," focuses on processes and analytics. This section illustrates why the manufacturing process plays such a large role in biotechnology, and to what extent product quality is determined by processes and analytics. Moreover, essential technologies for industrial manufacturing as well as methods and areas of application of analytical testing are described.

Part III, "Pharmacy," briefly elaborates on the basic principles of drug effects on humans and the essential steps of pre-clinical and clinical drug studies. The successful end of the clinical test marks the starting point of commercialization.

Product quality plays a crucial role in pharmaceutical manufacturing. Part IV, "Quality Assurance," elucidates the organizational and operative workflows for quality assurance, including the rules of GMP.

Almost all activities of commercial production happen in the framework of legal regulations. Part V, "Pharmaceutical Law," describes drug regulations and laws as well as institutions and enforcing official authorities.

The translation of process technology into large-scale manufacturing capacities is described in Part VI, "Production Facilities." Basic principles of the design of GMP-compliant manufacturing facilities are given and different building concepts compared. The planning process that leads to industrial plants is illustrated. Here, we include a brief look at the regulations regarding health, safety, environment, and construction that form the legal framework of industrial production facilities.

Commercial thinking is the spine of efficient production. Part VII, "Economy," introduces essential principles around product sales and cost of goods accounting. It compares concepts of in-house manufacturing with outsourcing strategies and elucidates the decision factors leading to capital investments in biotechnological plants.

The book closes with a Bibliography providing literature and web references and an appendix providing a list of abbreviations and an alphabetical index of keywords.

1.1.1 Relationship Between Production and Development

It is widely understood that production starts when development provides a marketable product and a commercially feasible manufacturing process. Ongoing market supply is secured by process optimization or the provision of additional manufacturing capacities, depending on how market demand develops. For

biotechnological pharmaceuticals, the flexibility to react to demand changes is reduced due to the following reasons.

Drug application and the manufacturing process are described and fixed in the regulatory license. As the biotechnological manufacturing process is a quality-determining factor, it has to be finally defined at the time point of regulatory submission and can thereafter be changed only with relatively high effort. The market application contains proof of the safety and efficacy of the drug; it adds to the complexity that in biotechnology, this proof has to be made – at least partly – with material from the commercial process and manufacturing site. Changes to the process or site require comparability exercises that can be more or less complex depending on the risk associated with the change. All of this means that the manufacturing process is fixed at a relatively early time point during development and can only be changed with quite some effort.

Clinical and process development

This coherence is illustrated in Figure 1.3. Product development consists of clinical development, on the one hand, and development of the manufacturing process and the analytical methods, on the other hand. The clinical development renders proof of safe and efficacious use of the drug in humans. Ideally, this proof is generated with material from the process and the site designated for commercial supply. There is a challenge with this ideal approach: if the process would be finally established and only after that clinical development be initiated, the timelines of development would add up unacceptably. Therefore, the different branches in the development workflow occur in parallel; different stages of the clinical development are supplied with different development stages of the manufacturing process.

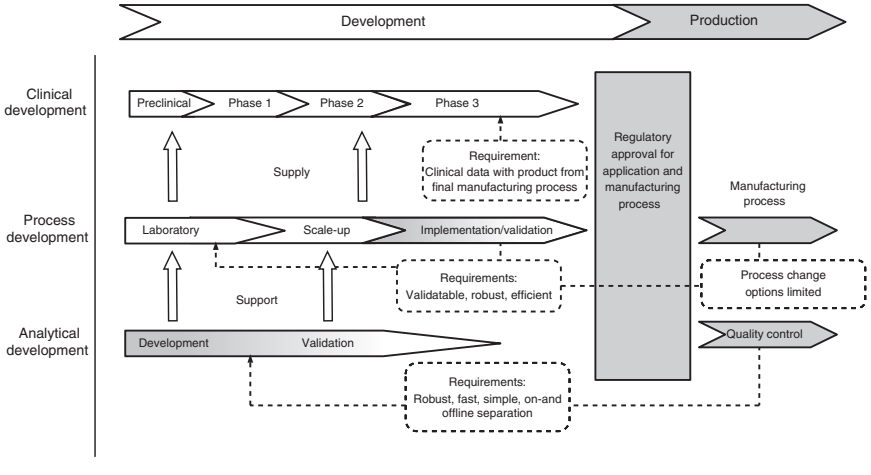

Figure 1.3 Relationship between production and development.

From lab to large-scale process

Coming from the laboratory scale, the process is evolved step by step into the final and mature manufacturing process. The scale and maturation a technical process achieves until first used for commercial supply depend on the characteristics of the process, on the requested product demand, and, often, on the available time for development. At the endpoint of development, the process is implemented in the designated commercial supply facility.

Validation and critical parameters

Product generated in these so-called full-scale runs must be used in representative amounts in the clinical trial. Process validation shows that the drug can be manufactured reproducibly and in good quality under consideration of applicable operating procedures. Product generated in these so-called "validation runs" must usually be used in representative amounts in the clinical trial.

An "easy-to-validate" process means that product quality is essentially independent of fluctuations of the critical process and equipment parameters. Critical parameters as well as measures to control them should be identified in the lab-scale process. These links result in interactions between production and development long before the actual supply to the commercial market.

Role of analytics

Owing to the heterogeneous composition of biological pharmaceuticals, analytical methods play a special role. Just like the process, the developed methods find their way into the regulatory license documentation. Concurrently – while being optimized itself – analytics has to support process development from very early on. Production requirements such as speed, robustness, and simplicity of testing methods have to be taken into account. Moreover, it has to be decided which method should support processing, which method is necessary for product characterization and process validation, and which method should be used only in the development phase.

This short outline illustrates how deeply the aspects of production reach into the development phase. An early recognition of production aspects can help to avoid detours and project delays.

Production facilities

An important interface, that is, not shown in Figure 1.3 is the one to the facility in which the manufacturing process is carried out. The capital investment in a manufacturing plant, and also the alternative contractual obligation with an outside source, means an additional financial risk of considerable size. This issue is further discussed in Part VII.

1.1.2 Relationship Between Production and Marketing

Production makes the final and packed product available for marketing (Figure 1.4). The packaging provides product protection and a possibility for attracting customers; especially in

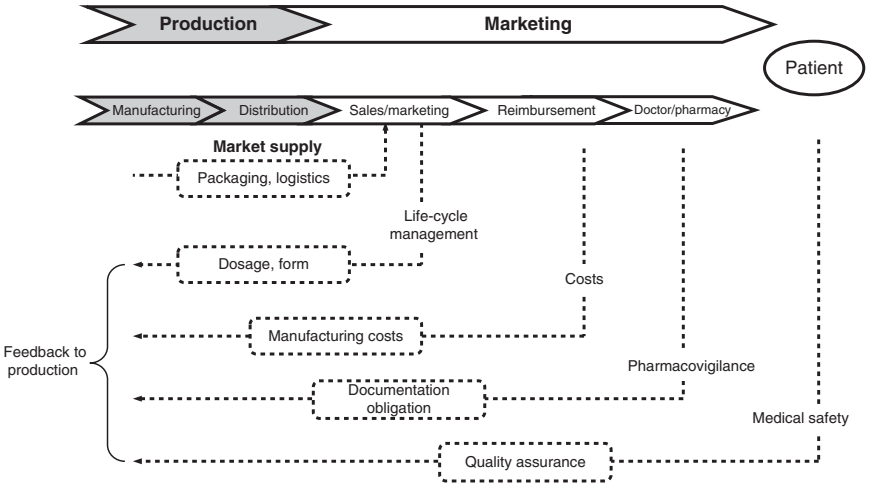

Figure 1.4 Relationship between production and marketing.

the pharmaceutical arena, the packaging contains a considerable amount of user information. The coordination and distribution of the country-specific final products is done by production logistics, which has to react flexibly to requirements from sales and marketing.

Life-cycle projects support the development of the project in the market place and usually affect not only the pharmaceutical dosage and form but also the indication of the product. In these cases, production has to adapt to changes in demand, packaging materials, or formulation processes.

Cost aspects

The acceptable manufacturing costs are determined by the achievable price in the marketplace, which is often regulated by country-specific reimbursement systems. The construction or maintenance of manufacturing plants has to be justified by adequate profitability calculations that are based on estimates and expectations of the market situation, and the desired profit margin, on the one hand, and the operating and capital expenses, on the other hand. These projections often reach far into the future (more than 10 years) and leave large room for variations.

Specialties of the pharma market

While the aforementioned characteristics also apply to other goods, there are indeed pharma-specific features, for example, the governmental monitoring system, the exceptionally high ethical responsibility of pharmaceutical companies, and the official regulation of drug reimbursement. Safety of patients is guaranteed by instruments used for pharmacovigilance and intensive product quality assurance. Pharmacovigilance systems serve to register unforeseen adverse effects of drugs and route them to the supervisory body. To achieve this goal, the pharmaceutical company

collects and evaluates blinded patient data; in case of an unforeseen adverse event, a root-cause analysis has to be performed. To perform this analysis, it is necessary that the specific medicament used by the patient can be traced back to the manufacturing site and batch. It is the specific batch documentation that then provides insight into whether deviations have occurred during the operation that might have influenced the quality of the product. If yes, it needs to be clarified in a second step whether such a quality variation could have triggered the adverse reaction. Thus, the requirements of pharmacovigilance lead to a comprehensive documentation obligation of the entire manufacturing process.

The target of pharmacovigilance is to recognize risks retrospectively. As a complement to the framework of drug safety, there are intensive measures for prospective quality assurance. This has a significant impact on the operational workflows as will be shown in Part IV.

1.2 Product Groups

Pharmaceutical biotechnological products can be classified into:

- Vaccines derived from non-genetically modified organisms or blood.
- Therapeutics from blood or animal organs (e.g., Factor VIII and insulin).
- Antibiotics manufactured traditionally in biological processes. Usually this is done with non-genetically modified organisms.
- Recombinant proteins (i.e., active ingredients) derived from cultivation of genetically modified cells. Including monoclonal antibodies, these represent the biggest sector of current pharmaceutical biotechnology.
- A new branch of therapy opens up with the possibilities of cell and gene therapeutics. These complex interventions into the human body require the reassessment of the pharmaceutical safety concept and demand special precautions from production technology and engineering.

Manufacturing technologies of different product groups can be similar

The focus of the present work lies in the production of therapeutic recombinant proteins; however, the principles described can be applied to the other product groups as well. A closer look at the groups reveals interesting therapeutic and technological overlaps. For example, innovative gene therapy can learn from experiences in virus production gathered in the conventional vaccine field. Also, vaccines will face a new era due to the possibility to produce

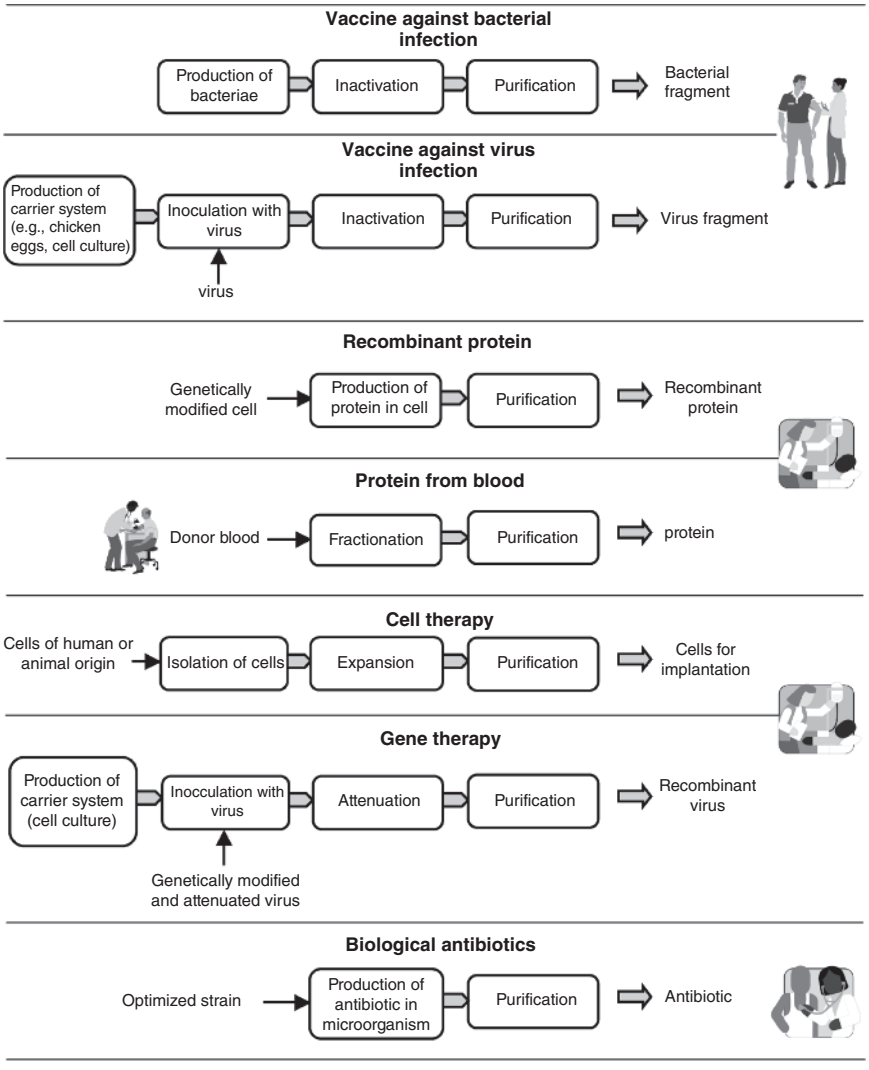

Figure 1.5 Schematic production workflows of important product groups. Product groups are shown on the right. Attenuation = elimination of reproducibility, but retention of infectivity; inactivation = elimination of reproducibility and infectivity.

monoclonal antibodies (Section 1.3.2.1). In the following text, the product groups – with the exception of antibiotics – will be covered in more depth.

Figure 1.5 schematically shows the production workflows for different product groups. There are differences regarding the genetic modification of the starting material. Genetically modified

organisms are mainly deployed for recombinant proteins and gene therapeutics, but cell therapy can also use this technology. The products can be proteins, viruses, bacterial fragments, cells, or intact viruses for gene therapy.

1.2.1 Vaccines

There are two principles of vaccination:

- *Passive vaccination*: Antibodies against the pathogen are administered.
- *Active vaccination*: The immune system is confronted with alleviated pathogens and builds up its own immune defense against the causative organism.

Antibodies for passive vaccination are prepared by injecting the pathogen into animals. The immune system of the animals pours out so-called polyclonal antibodies into the blood system. Blood is collected from the animals, and the antibodies isolated and purified, so that they can be administered to humans.

Active vaccination uses inactivated germs that are no longer pathogenic, but still immunogenic. Activation allows the immune system to recognize the real pathogenic germs much faster and therefore fight them before they can spread out and cause the illness. It suffices to present only a moiety instead of the whole pathogen to enable the immune system to recognize the substance foreign to the body. This moiety can be the hull protein of a virus, whole inactivated cells, or pathogen-specific deoxyribonucleic acid (DNA). The general term for these immune response-inducing agents is "antigen." Active vaccines like the influenza vaccine can be proliferated in chicken eggs and reworked to vaccines.

1.2.2 Pharmaceuticals from Blood and Organs

Many diseases can be attributed to the lack of certain proteins in the blood. In part, these proteins can be extracted from animal or human blood or organs, such as insulin against diabetes or Factor VIII against bleeding disorders. Biotechnology has made it possible for these proteins to be obtained without being tied to these expensive and – under aspects of safety – questionable raw materials from natural sources. In some cases, blood-derived products still play a role as it has not yet been possible to successfully replace them completely by recombinant proteins.

Plasma fractionation

To isolate the proteins from the blood, it is first separated into its two main components: plasma and cells. The plasma is further fractionated to obtain the proteins. It is associated with considerable analytical and organizational effort to guarantee the safety of the raw material blood, especially the absence of viral contamination and transmissible spongiform encephalopathy (TSE)-inducing components. Despite the intensive surveillance of blood donors, the danger of safety-relevant incidents persists. It can be expected that the production of proteins will be more and more shifted to recombinant technologies, while whole-blood donations will remain irreplaceable for patient treatment in hospitals.

Risks of protein extraction from blood

1.2.3 Recombinant Therapeutic Proteins

Recombinant proteins, including monoclonal antibodies, by far, make up the largest group of biotechnological pharmaceutical products. Table 1.1 shows some examples; in addition to the medical indication and the functionality in the human organism, it provides details regarding the size and type of the molecule. Section 1.3 gives further insight into the structure of proteins and the terms of amino acids and glycosylation.

A huge growth potential is expected for monoclonal antibodies and antibody fragments.

The starting point for all protein production is the genetic modification of the host cell in which the protein should be expressed. The endpoint usually is a parenterally (per injection) administered product in liquid or solid form.

1.2.4 Cell and Gene Therapeutics

***Cell therapy**: implantation of intact cells*

Therapeutic proteins are administered to compensate for the lack of the respective natural protein in the organism. As the molecule is eliminated either by degradation or excretion, the administration has to be repeated to achieve a constant active agent level. In contrast, cell and gene therapy is based on the idea of fighting the disease at its source and enabling body cells to express the missing proteins by themselves. Figure 1.6 schematically illustrates the differences between the philosophies of protein versus cell and gene therapy treatment. The starting point is a disease caused by a lack of the example protein X. In conventional therapy, the protein is produced *ex vivo* and injected into the patient. Owing to elimination processes, the protein disappears after a while. In gene therapy,

Table 1.1 Examples for recombinant proteins.

Name	Indication	Functional group	Number of amino acids; glycosylation, and fraction of sugars of molecular weight; molecular weight
Insulin	Diabetes	Hormone	AA 51; Gly no; 5.8 kDa
Human growth hormone	Dwarfism	Hormone	AA 191; Gly no; 22.1 kDa
Factor VIII	Bleeding disorder	Clotting factor	AA 2332; Gly to 35%; 300 kDa
Lepirudin	Thrombosis	Anticoagulant	AA 64; Gly no; 7 kDa
Tissue plasminogen activator	Thrombosis	Thrombolytic agent	AA 72; Gly to 25%; 72 kDa
Interferons (IFNs)	Diverse: multiple sclerosis, hepatitis, arthritis, and so on	Immune modulator	IFN-β: AA 166; Gly yes and no; 18.5 kDa
Interleukins (ILs) (13 different types)	Diverse: asthma, HIV, cancer, mucositis, and so on	Immune modulator, signal agent between immune cells	IL-2: AA 133; Gly yes; 15.5 kDa
Erythropoietin	Anemia	Growth factor	AA 165; Gly to 40%; 34 kDa
Granulocyte colony-stimulating factor (G-CSF), granulocyte macrophage colony-stimulating factor (GM-CSF)	Infections, cancer	Growth factor	G-CSF: AA 174–180; Gly yes; 19.6 kDa; GM-CSF: AA 127; Gly yes, 15.5, 16.8; and 19.5 kDa
Monoclonal antibodies	Cancer, transplantation, and so on	Antibodies	IgG: AA about 1300; Gly yes; about 150 kDa

Gly = glycosylation; AA = amino acids; Gly to 30% = molecular weight fraction of glycosylation can reach up to 30%.

a genetic sequence is introduced into the body, which contains the construction plan for the desired protein as well as the capability to infect suitable target cells. This combined capability is generated by means of biotechnological methods: the protein-encoding gene is linked to a molecular "ferry" (or vector) that carries the gene into the designated target cells. This ferry is a virus that has been modified in such a way that it retains its infectivity but lacks its ability to replicate. After having been infected the cells start to produce the desired protein. In the ideal case – if the construct is genetically very stable and the expression rate high – this process has to be carried out only once. If the modified viruses are injected directly into the body, it is called *in vivo gene therapy*, but it is also possible that the cells are extracted from the patient and re-implanted after being infected in the lab (*ex vivo* gene therapy). The latter is basically a crossover between cell and gene therapy.

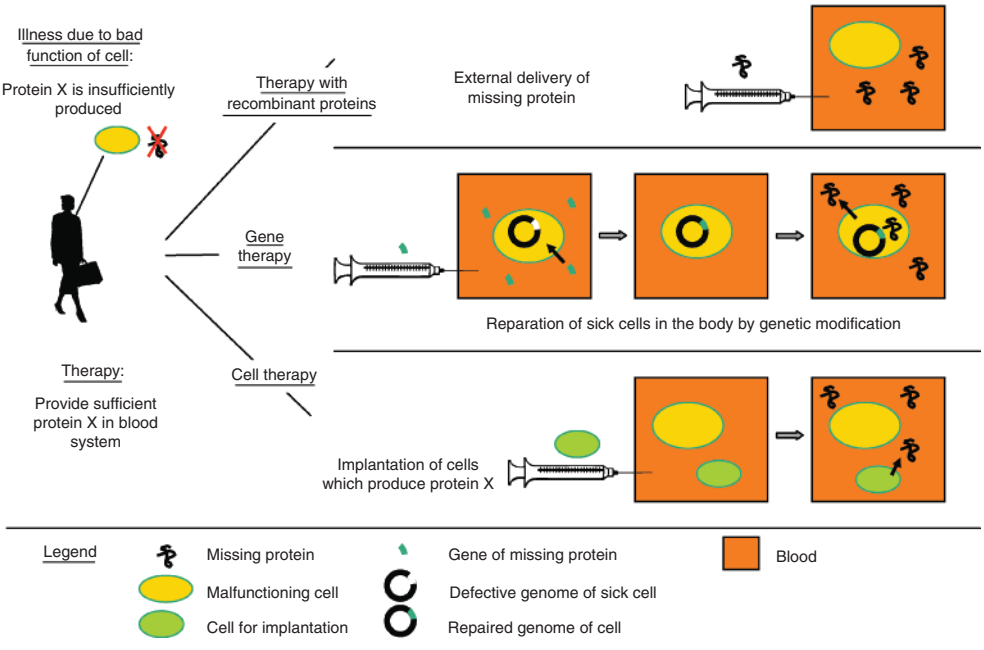

Figure 1.6 Schematic comparison of traditional therapy and cell and gene therapies.

The basic principle of cell therapy is to convey cells to the body that have the desired functionality; thus, cell therapy does not aim at repairing dysfunctional cells, but rather replacing them. These cells can originate from animals (xenogenic) or humans, either patient proprietary (autologous) or non-patient proprietary (allogeneic). The allogeneic and xenogenic approaches raise questions regarding immunogenic responses, comparable to tissue rejection in organ transplantation. If the cells are genetically modified, they belong to transformed cell lines; cells that have not experienced any genetic modifications are called *primary cells*. Thus, "allogeneic cell therapy with primary cells" denotes a therapy in which cells of a foreign donor are implanted without genetic modification into the patient. The functionality of cells is not restricted to protein production – other molecules can also be expressed. A prominent area of cell therapy is the replacement of tissue or organs (tissue engineering). Consequently, the widely discussed therapy with human stem cells is only one form of cell therapy, which is characterized by the adoption of non-differentiated stem cells.

Cell production for therapy

Cell therapy production starts with isolating the cells intended for implantation. This is clearly a crucial issue of cell therapy, as the starting material is limited and the subsequent expansion of the

cells is restricted by the natural limit of generation numbers. The implantable cells are available for surgical implantation after an additional manipulation step such as washing or buffer exchange.

Gene production for therapy

The genetic construct for gene therapy is produced by proliferating the DNA in a host organism that can be first expanded and subsequently transfected with a modified virus. This infected cell produces the desired virus. To control infectivity, the virus can be weakened (attenuated); at the end of this production process, the product consists of the attenuated, modified virus.

Safety questions around cell and therapy

Cell and gene therapies are at the advent of their development. Despite the fact that the approaches seem plausible, complex questions around drug safety arise. The administration of generally replicable and propagatable substances is very different compared to conventional protein therapy. The scarce source and the handling of the living "cell" system impose new challenges on production and distribution processes.

1.2.5 Antibiotics

Antibiotics have been produced biologically for many decades. A penicillin-producing yeast strain is cultivated in a biological fermentation step, and the expressed penicillin is further purified into a pharmaceutical product. In particular, fermentation resembles that with recombinant microbial expression systems (Section 2.3.1). In contrast, purification is different because penicillin is a relatively small and robust molecule compared to proteins.

1.3 Basics of Biology

This section is dedicated to some basic principles of biology and biochemistry that are relevant for the understanding of production processes on the level of this book. After a short outline of cell biology and microbiology, the four basic molecular entities of biochemistry are introduced: proteins, nucleic acids, polysaccharides, and lipids.

1.3.1 Cells and Microorganisms

Each form of life – plants, animals, or microorganisms – consists of biological cells. While plants and animals constitute themselves

as enormous networks of different cell types, microorganisms are predominantly single celled.

Microorganisms can survive in their selected habitat independently, while cells of higher organisms depend on their united cell structure. Cells are characterized by some general features:

- Cells contain a carrier of genetic information (double-helical DNA) and a single-strand ribonucleic acid (RNA) derived thereof. During replication – in the process of propagation – its genetic information is prone to erratic variations (mutation).
- Cells exchange nutrients and waste products with the environment (metabolism) for the purpose of energy recovery (catabolism) or substance construction (anabolism).
- They are confined by a membrane that allows for controlled substance exchange with the environment.
- They communicate via so-called receptors with the environment and can react to changes in external conditions.
- They are capable of replicating themselves and a number of higher cells differentiate into other cell types.

In addition to genotypic characterization of cells and organisms, based on gene technology, there is a phenotypic characterization based on the differences in shape, movement pattern, staining (Gram staining), metabolic pattern, and preferred habitat.

Mycoplasmas are special bacteria. They do not possess a cell wall and are exceptionally small. Obviously, they are resistant against types of antibiotics that attack cell walls. As with viruses, they are not retained by filters of a pore size of 0.22 µm. Some members of this family are pathogenic (e.g., *Mycoplasma pneumoniae*, *Mycoplasma genitalium*).

Viruses are not cells but consist of encapsulated genetic information in the form of DNA or RNA. They do not have their own metabolism and depend on other biological cells for their replication. Many representatives of this group are pathogenic [e.g., human immunodeficiency virus (HIV), hepatitis, and herpes].

1.3.1.1 Structure and Types of Cells

There are two types of cells: the simple prokaryotic cells and the more complicated eukaryotic cells.

Chromosome and plasmids

Structure of Prokaryotic Cells Prokaryotic cells consist of a cytoplasm surrounded by a membrane, which is itself surrounded by a stabilizing cell wall. The most important functional units embedded in the cytoplasm are the ribosomes, the chromosome, and the plasmids (Figure 1.7).

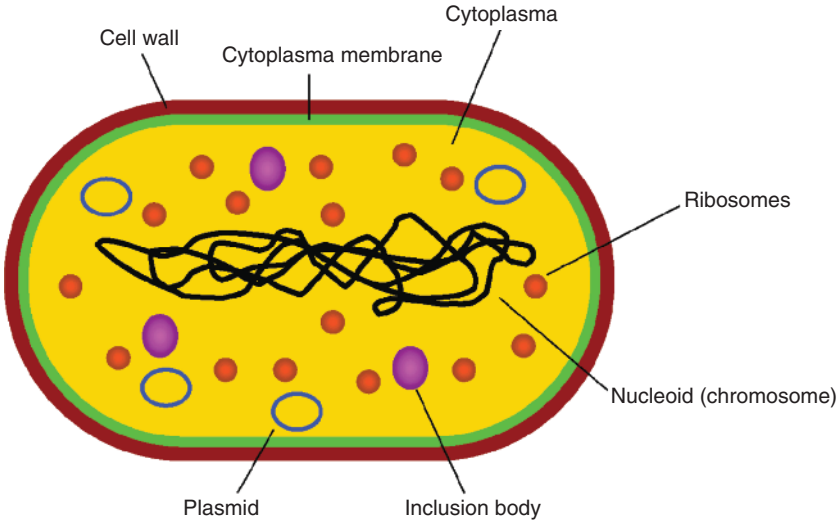

Figure 1.7 Schematic of a prokaryotic cell.

Inclusion bodies and secretion

The chromosome is a single-stranded DNA double helix and contains the genetic information for the construction and replication of the cell. Protein biosynthesis from the DNA happens at the ribosomes after the DNA information has been transcribed into RNA. In addition to the chromosome, prokaryotes often carry further genetic information in the so-called plasmids. These are ring-shaped DNA molecules located in the cytoplasm. They usually encode secondary functional proteins such as the substances enabling penicillin resistance. The inclusion bodies shown in Figure 1.7 are storage locations for substances that for the time being are not required.

Gene expression

The contemplated functional units are highly relevant in biotechnological production. Genetic modification can either be made in the plasmids (also denoted as episomal or extrachromosomal) or in the actual chromosome (also denoted as integrative or intrachromosomal). The modification consists of the insertion of the protein-encoding DNA sequence into the plasmid or chromosome, which ultimately leads to protein production at the ribosomes (Section 1.3.1.5). This so-called gene expression can be controlled by external factors like temperature or supply of a specific agent (induced promoter) or happen spontaneously as a natural cellular activity (constitutive promoter). If the proteins are delivered through the cell membrane into the environment, this is called "secretion," but proteins are often aggregated in inclusion bodies (Section 1.3.1.1).

Inclusion bodies

Structure of Eukaryotic Cells Compared to prokaryotic cells, eukaryotic cells are strongly compartmentalized into functional units (organelles). As they do not exist alone, but in a network of other cells, they have – unlike prokaryotic cells – no cell wall but a structurally relatively weak cytoplasmic membrane. Unlike prokaryotic cells, the chromosome is contained in a cell nucleus. The genetic information of eukaryotic cells is much more complex than that of prokaryotic cells (Figure 1.8).

The ribosomes, which are responsible for the synthesis of proteins, are located outside the nucleus in the cytoplasm. Proteins are transferred from there into the endoplasmic reticulum for further modifications. The mitochondria are the power plants of the cells; they generate the chemical energy to support all cell activities, such as molecule synthesis or directed transport of substances. The dictiosomes are the "glands" of the cell. Their role is to prepare secretion of agents from the cell. Therapeutic proteins are expressed either via intrachromosomal integration of a desired DNA sequence or implantation of plasmids into the cytoplasm. Eukaryotic cells are approximately 10 times larger than prokaryotic cells.

1.3.1.2 Metabolism

Cells constantly exchange substances with their environment. This activity is called *metabolism*; it serves to generate energy to support the living system (catabolism) and to absorb substances that are needed for creating new cell material during, for example, cell separation (anabolism).

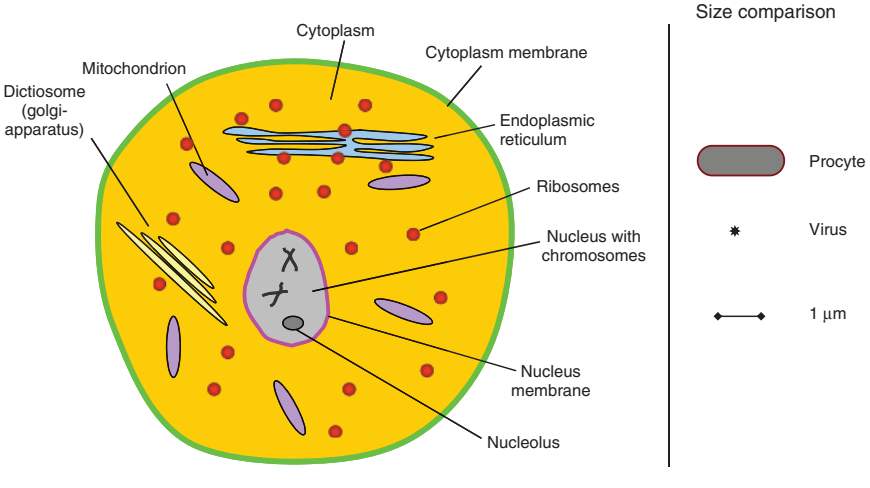

Figure 1.8 Schematic of a eukaryotic cell.

Catabolism

Energy is generated by degrading carbohydrates, transferring them from a high to a low chemical energy level (Figure 1.9). Under addition of oxygen, the carbohydrate molecule converts its hydrogen atoms into water and is finally decomposed to carbon dioxide. This oxidation reaction resembles an incineration with the exception that the cell controls the released energy and – along the reaction cascade – stores it in small usable portions, namely, the molecule adenosine triphosphate (ATP).

In the absence of oxygen, the carbohydrate cannot be decomposed completely. Unfinished oxidation is called "glycolysis"; in this case, the cell lives under anaerobic conditions (anaerobic = without oxygen). However, most biopharmaceutically relevant organisms depend on aerobic respiration, which means on the supply of elementary oxygen being dissolved in the ambient water. In case of reduced oxygen supply, the organisms can reduce their metabolism and produce other substances than if they had sufficient oxygen. That is, the main reason why homogeneous aeration of bioreactors is of great importance.

Synthesis metabolism

During anabolism (synthesis metabolism), the energy recovered in catabolism is used to create the molecular building blocks of the cell (amino acids, purine, and pyrimidine bases, sugar phosphates, organic acids, etc.) from the available nutrients (e.g., glucose, fatty acids, trace elements). The building blocks are utilized to assemble the biological macromolecules constituting the chemical backbone

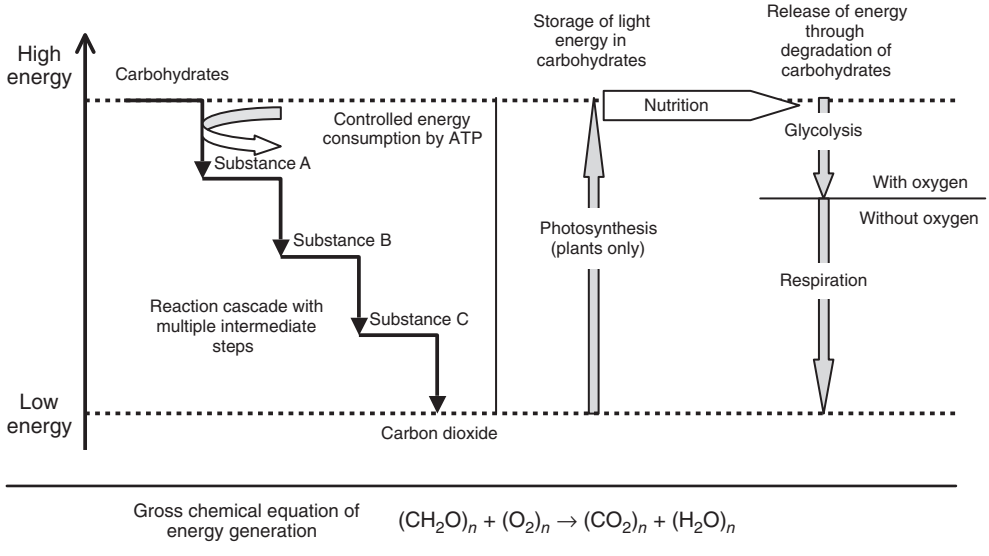

Figure 1.9 Schematic of energy metabolism.

of biological life (nucleic acids, proteins, lipids, polysaccharides; Section 1.3.2). The metabolic pathways consist of highly complex chemical reactions including protein synthesis. The construction plans for all proteins are laid down in the DNA. After being activated, this information is transferred to the ribosomes where the actual synthesis takes place.

Nutrient media

In industrial fermentation, nutrients have to be fed with the culture medium. Here, microorganisms are much less demanding than animal cells. While the former just require a menu of fundamental nutrients (carbon source, trace elements), the latter – to grow – depend on the provision of "blood-like" conditions (vitamins, hormones, amino acids), resembling their original habitat.

1.3.1.3 Reproduction and Aging

Time for cell division

Cells proliferate by division, resulting in an identical copy of the original cell. The time that a cell needs to divide is determined by the complexity of the chemical reactions needed to synthesize the elements of the new cell. That is, why the doubling time of simple prokaryotic cells is much faster than that of the much bigger eukaryotic cells. Under optimal growth conditions, the bacterium *Escherichia coli* needs 15–20 min for division, the yeast *Saccharomyces cerevisiae* 2 h, while mammalian cells have doubling times of up to 24 h.

Cell growth

Provided that sufficient nutrients are available, microorganisms divide spontaneously and permanently as illustrated in Figure 1.10. Mammalian cells usually need an external trigger for growth. They often form adherent monolayers (i.e., they exclusively grow in single layers bound to surfaces), are contact inhibited, and cease to grow once the surface is covered entirely. The total number of divisions is limited for most cell types. To achieve an unlimited growth potential, so-called continuous cell lines have been developed. This was achieved by hybridizing the genetic information from a tumor cell and a different desired cell line.

Genetic stability

Cells die after reaching their species-specific lifespan (apoptosis) and are prone to mutation. This spontaneous variation of genetic

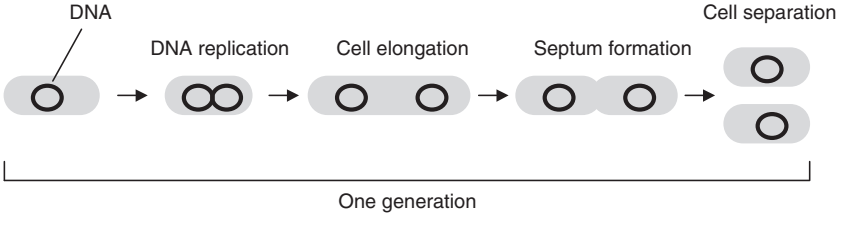

Figure 1.10 Schematic of cell division.

information can be caused by transmission errors during DNA replication as well as other factors. The genetic stability (i.e., the transfer performance of genetic information from generation to generation) is an important criterion for the evaluation of a modified cell line.

Section 2.4 will provide further details on culture doubling times and Section 2.3.1 will elaborate on characteristics of different cell types.

1.3.1.4 Viruses and Bacteriophages

Viruses and phages are not cells

Viruses and phages are not constructed as cells but consist of genetic information in the form of DNA or RNA, that is, encapsulated in a simple coat. They do not have their own metabolism and depend on cells for proliferation. This goal is achieved by transferring their genetic code into the cells (infection). While viruses affect animal cells, phages affect bacteria.

Many viruses are known to be pathogenic. The principle of active vaccination against virus-related diseases (smallpox, influenza) is that inactivated (non-augmentable) viruses are injected to evoke an immune response.

Viruses and phages can transfer genetic information

The ability of viruses and phages to transfer genetic information to cells has a great significance in biotechnology. A negative aspect is that the risk of viral contamination jeopardizes the safety of the drug; on the positive side, gene technology uses the ability of viruses and phages to introduce genetically modified DNA strands into cells.

Viral contamination

The risk of viral contamination generally exists for bacterial and mammalian cells. Owing to the inability of bacteriophages to infect human cells, and the inability of human viruses to infect bacterial cells, the virus threat for the patient is higher in mammalian cell culture than in microbial production processes. Figure 1.11 illustrates the main pathways of virus access to the product. There are several processing aids, recovered from human or animal blood, which have to be tested very diligently for the absence of viruses (HIV, hepatitis, etc.). Improper handling (e.g., in the lab when generating the cellular production system; Section 2.3.2) can lead to infection of the cell material. Particularly in this early stage of production, viral contamination can remain undiscovered as it may not lead to a detectable change of the cells.

In the simplest case, the viruses proliferate in the cells outside of the host genome, destroy the cell, and spread out large quantities of viral particles to the environment. Unfortunately, some viruses and phages integrate into the cell genome, are passed on over generations, and are only activated by specific external trigger mechanisms. Others reside in the cell and let the cell produce and

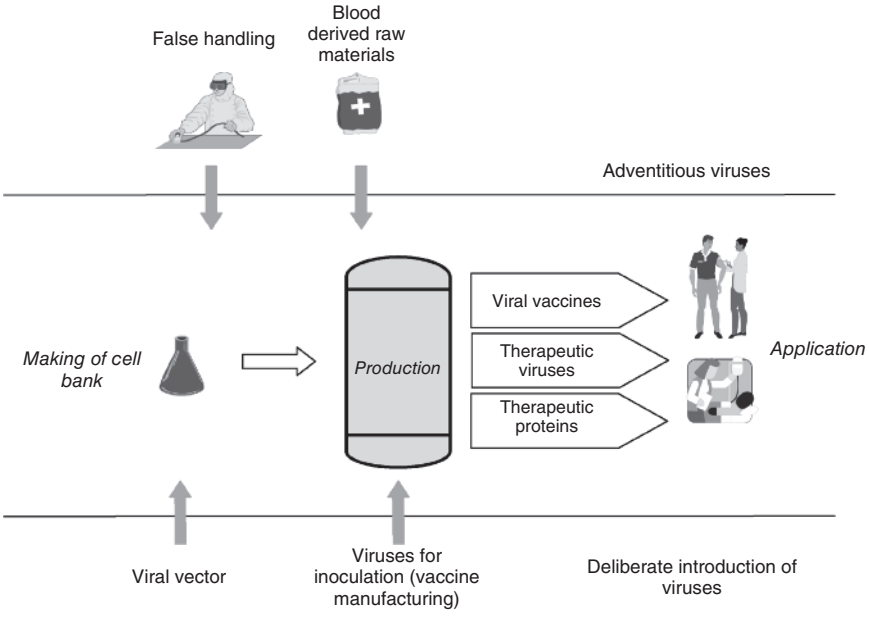

Figure 1.11 Intentional and unintentional addition of viruses (or phages) in the general process scheme. Upper = unintentional contamination; lower = intended functionality.

secrete viral successors continuously. In both cases, the contamination can remain undetected for an extended period of time.

In biotechnology, the ability of viruses to inject DNA into biological cells is used for genetic manipulation. Viruses function as "vectors" to introduce genetic information into the genome of cells, which has been implanted beforehand into the virus (transfection). For that purpose, the virus is deprived of its pathogenic and cell-destructive effects. Transfection can be used to generate a host cell line for protein production, to inoculate a culture for vaccine production or for direct injection into humans for the purpose of gene therapy (Section 1.2.4).

1.3.1.5 Protein Biosynthesis

The vast majority of therapeutically active substances in biotechnology are proteins. How can these proteins, which naturally are generated in the human body, be produced outside man?

The answer is as simple as it is surprising: nature has laid down the building plans in the form of the genetic code. This code, which is present in any biological cell as a DNA molecule, is universal – a certain sequence of links in the molecular chain, be it in a bacteria, plant, or human cell, leads to the synthesis of essentially the same protein molecule. If the human genetic code for a protein

can be introduced into a biological cell, this cell produces the desired protein, which can later be separated and administered as a therapeutic agent.

The important mechanisms of protein biosynthesis are illustrated in Figure 1.12. DNA consists of four molecular elements, the so-called bases: adenine (A), guanine (G), thymine (T), and cytosine (C). The whole genome is composed of these four molecules. A combination of three bases, called *a triplet or codon*, encodes for one amino acid (i.e., a sequence of codons encodes a certain sequence of amino acids and thus a protein). There are also signal codons that, for example, mark the beginning and the end of a protein assembly.

Transcription and translation

Protein synthesis starts with the transcription of the DNA into RNA, which is unique since each base of the DNA has a corresponding complementary base in the RNA molecule. The RNA transfers to the ribosomes, where it is decoded and delivers the information for the protein sequence (translation). After being synthesized at the ribosomes, the proteins then assume their role in and outside the cell.

Gene expression

Thus, the goal of genetic manipulation is to introduce the genetic code for the protein into the cell and let the ribosomes produce the amino acid chain. This realization of genetic information into functional protein molecules is called "gene expression" and the intensity

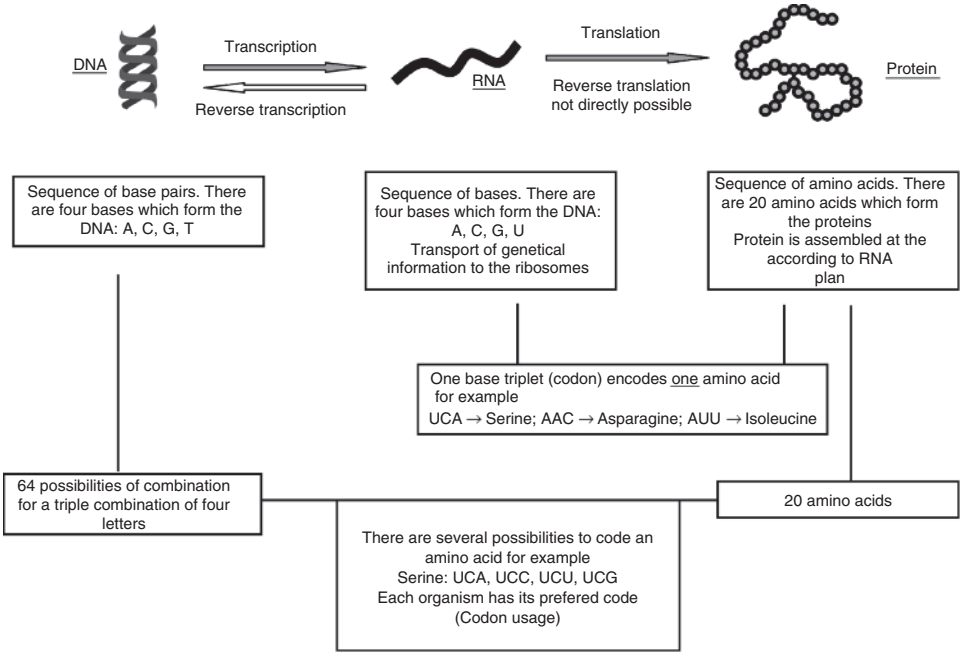

Figure 1.12 Basic principles of protein biosynthesis.

of the production is called "expression strength." The gene can either be integrated into the chromosome or exist extrachromosomally as a plasmid. As there are 64 possible triplets, but only 20 amino acids, some triplets encode for one and the same amino acid. It appears that different organisms have different preferences for the genetic code. The expression strength is higher if one code is selected instead of the other. This "genetic dialect" is called "codon usage."

Post-translational modifications

Regardless of which type of cell (plant, bacteria, human, or animal) is chosen, the sequence of amino acids in the protein can be generated anywhere; however, the functionality of many therapeutic proteins not only depends on the amino acid sequence but also to a great extent on the so-called post-translational modifications (Section 1.3.2.1). Usually these modifications can only be realized in eukaryotic cells (yeasts, plants, animals, humans).

1.3.2 The Four Molecular Building Blocks of Biochemistry

Composition of cell

The technological and therapeutic approaches of biotechnology require knowledge of the basic fundamentals of biochemistry. After all, the fascinating functionality of biological systems is founded on biochemical processes. A typical cell of the simple bacterium *E. coli* consists of about 70% water; the rest is called the *dry weight*: 55% of the dry weight is a subset of the 2500 known proteins, 20% is RNA and 3% DNA, and 17% polysaccharides and lipids. Approximately 5% is not macromolecules, but substances of low molecular weight such as ions, and precursors of amino acids, sugars, and nucleotides.

1.3.2.1 Proteins

Proteins function in the body as enzymes, transport proteins, receptor proteins for cellular communication, messenger (hormones), defense agents (antibodies), contractible fibers in muscles, and structural molecules in bones, cartilage, and connective tissue.

Amino acids, peptides, and proteins

This broad functionality is provided by combinations of only 20 amino acids that are lined up in the protein as a linear chain. Up to 2300 amino acid building blocks are tied together by chemical peptide bonds. These bonds are located between the amino end and the carboxyl end of two amino acids. The beginning of the chain is called the *N-terminus* and the end is called the *C-terminus*. Shorter amino acid molecules of up to 50 amino acids are also denoted as peptides or peptide chains.

Protein folding

The linear chain folds itself into a three-dimensional structure, that is, essential for the biological functionality of the protein.

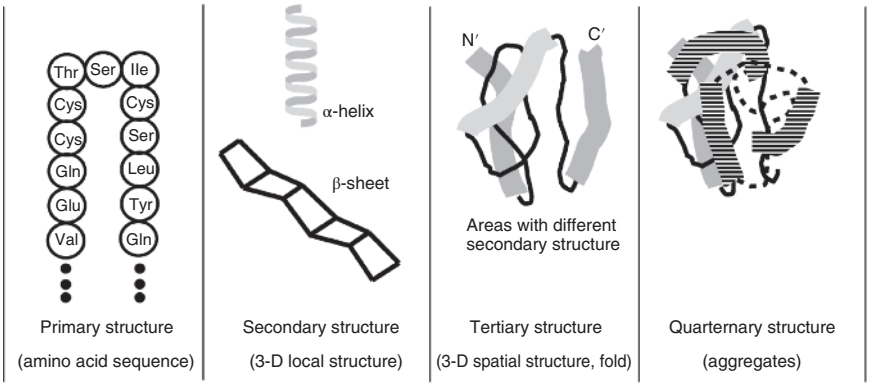

Figure 1.13 Protein structures.

The structures are illustrated in Figure 1.13. Folding is dominated by the formation of disulfide bridges between cysteine molecules, other post-translational modifications, and ambient conditions (pH value, salt content, temperature). As the term indicates, "post-translational modifications" are changes that are made to the protein after translation. They can have decisive influence on biological activity, stability, solubility, plasma half-life, and immunogenicity of the protein. There are several types of post-translational modifications; the most important are disulfide bridges and the attachment of sugar molecules to the primary structure (glycosylation). Sugar molecules can make up to 45% of the molecular weight of the protein (Table 1.1). Proteins with similar primary and secondary structure, but different tertiary structure, usually display completely different functionalities. The grouping of equal molecules – denoted as aggregate formation – can also have an effect on therapeutic biological activity.

Aggregation

Denaturation

Unfolding of the protein is also referred to as "denaturation," meaning the destruction of the three-dimensional structure, and typically goes along with a loss of biological activity. As long as primary and secondary structures are intact, the denaturation is usually reversible.

Figure 1.14 shows a typical example for a non-glycosylated protein – human insulin. The protein is composed of two chains linked by disulfide bridges at the cysteine amino acid molecules. The A-chain has 21 amino acids and the B-chain has 30 amino acids. With a molecular weight of 5.7 kDa, it is one of the smaller proteins.

An example of the schematic structure of one of the biggest serum proteins – immunoglobulin G (antibody IgG) – is shown in Figure 1.15. The molecular weight lies between 146 and 165 kDa

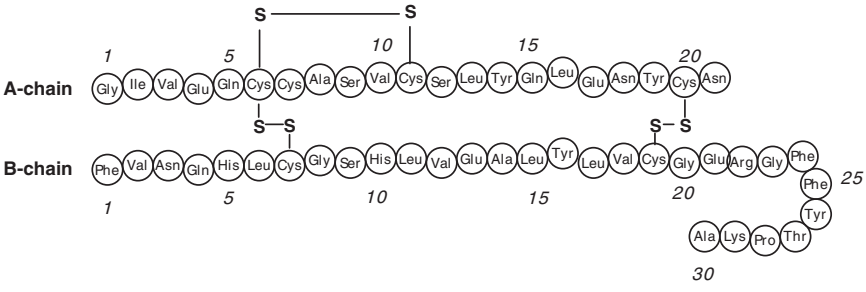

Figure 1.14 Molecular structure of human insulin. S–S disulfide bridges between cysteine molecules.

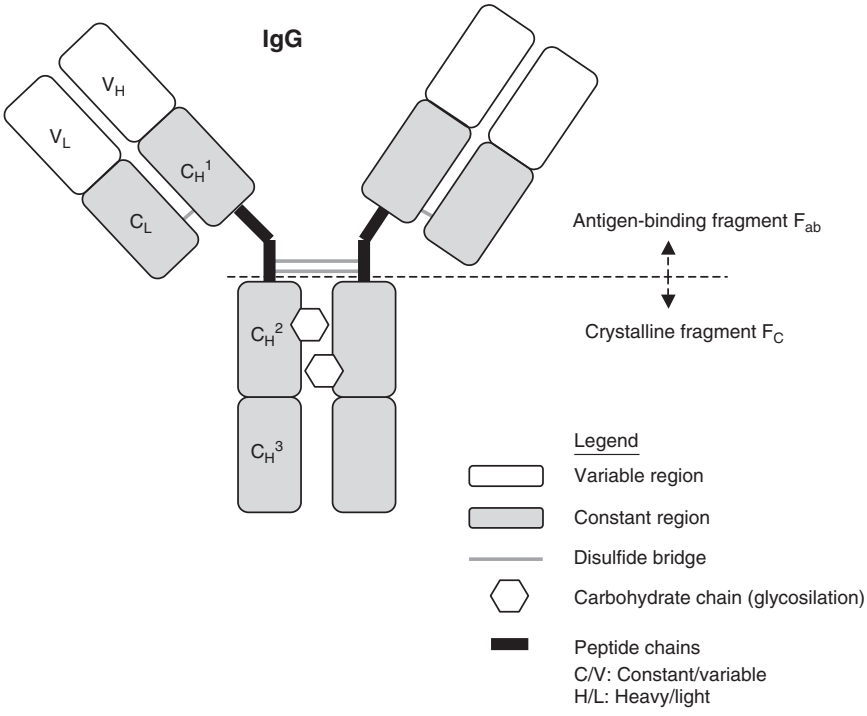

Figure 1.15 Schematic of antibody IgG.

depending on the molecule type. The molecule has carbohydrate chains (glycosylation) that are essential for the stabilization of the structure.

The boxes in Figure 1.15 indicate folded protein structures. The antibody consists of a so-called crystalline fragment (Fc, bottom) and an *antigen-binding fragment* (*Fab*, top). The variable regions are located in the upper part of the Fab, which are adapted to the foreign protein to be attacked. Both fragments are connected by

protein chains. In the vertical orientation, the antibody is symmetrical and connected by disulfide bridges. The molecule consists of two so-called heavy and two light chains that are connected by disulfide bonds. If antibodies are manufactured from one defined biological cell line, they are called *monoclonal antibodies*. In contrast, polyclonal antibodies are recovered from immunized animals.

To understand biological production processes and analytical methods after this short introduction into the structure and function of proteins, it is important to become familiar with some of the physicochemical properties of proteins.

Amino acids determine the physical properties of the protein

Individual amino acids have special features that determine protein solubility in water or organic solvents and their interaction with solid surfaces. Both properties are important for production and medical applications. Neutrally charged, water-adverse (hydrophobic) areas in the molecule are formed by the amino acids glycine, alanine, valine, leucine, and isoleucine. Polar, water-friendly (hydrophilic) structures result from the acidic amino acids asparagine acid, asparagine, glutamine acid, and glutamine, and the alkaline amino acids lysine and arginine. The other amino acids exhibit special functional groups or heterocyclic moieties.

Proteins are sensitive to thermal and mechanical stress

Proteins are unstable with regard to temperature, mechanical shear, and chemical stress (e.g., contact with oxygen) and are sometimes light sensitive. This does affect production technologies, sterilization options, storage, and transport conditions. The types of instabilities can be classified as follows:

- *Physical instabilities*
  - *Adsorption*: Protein attaches to container material. This results in a decrease of dissolved protein concentration or film formation.
  - *Denaturation*: Destruction of three-dimensional, primarily tertiary structure.
  - *Aggregation*: Formation of protein aggregates that not only can have a negative impact on drug efficacy but also can be toxic.
  - *Precipitation of protein from the solution*: Happens by aggregate formation or supersaturation, for example, due to temperature shifts.
  - *Association*: Chemical bonding between proteins that can lead to aggregation and precipitation.
- *Chemical instabilities*
  - *Oxidation*: Separation of hydrogen atoms and attachment of an oxidizing agent (methionine, histidine, tryptophan, cysteine).

— *Hydrolysis*: Clipping of peptide bond.
— *Deamidation*: Separation of the amide group (e.g., from asparagine or glutamine).
— *Disulfide-bridge clipping* (cysteine).
— *Beta-elimination*: Clipping of other functional groups of the amino acids.

Amphoteric behavior

Proteins show amphoteric behavior – their total charge is influenced by the pH of their environment. This forms the basis for the selectivity of several protein separation processes. Figure 1.16 illustrates this behavior. A protein in acid bulk phase binds to a cation exchanger; the same protein in alkaline conditions binds to an anion exchanger. The pH value where positive and negative partial charges equalize each other is denoted as the "isoelectric point" (pI). Under these conditions, the protein does not bind to polar matrices; hence, its solubility in water is minimal at the pI, limiting the freedom to change the pH value.

1.3.2.2 Nucleic Acids

The genetic information of the cell is stored in the ribonucleic acids DNA and RNA. The DNA constitutes the genome of the cell and the RNA is responsible for translating information stored in the DNA into functional proteins.

The DNA is constructed as a chain of nucleotides, which on their part are composed of three elements, out of which the most important is the base. There are four bases – and consequently four

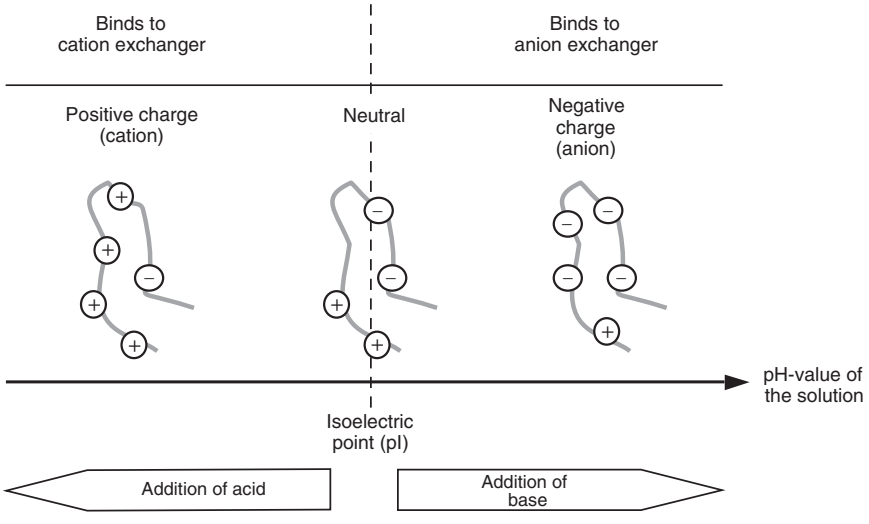

Figure 1.16 Amphoteric behavior of proteins. The ambient pH value determines the charge of the protein.

nucleotides – which make up the DNA: adenine (abbreviated A), guanine (G), thymine (T), and cytosine (C). The RNA has an additional base: uracil (U). DNA is structured as a helical double strand in which complementary bases are located opposite to each other. Its size is declared in numbers of complementary base pairs. RNA is a single nucleotide strand without paired or helical arrangement.

The genetic information can be displayed as a sequence of bases. Each amino acid is encoded by a combination of three subsequent bases (base triplet); for example, UCA stands for serine.

Genome

The entirety of the genetic information of an organism is called "genome." Each cell of an organism contains the complete genome of the organism. The genome of a bacterium is a simple ring-shaped DNA molecule; higher organisms have distributed their genome over several chromosomes each of which is a lengthy DNA molecule. For example, the human genome holds 23 chromosomes with a total number of approximately 3 billion base pairs. The genome of the protozoa *E. coli* has 4.6 million base pairs. At cell division, the genome has to be replicated – this lasts approximately 20 min for *E. coli*, while higher organisms need much more time.

1.3.2.3 Polysaccharides

Polysaccharides are molecular chains composed of sugars. Sugars (carbohydrates) are molecules that contain the elements carbon, hydrogen, and oxygen in the ratio 1 : 2 : 1 (e.g., glucose $C_6H_{12}O_6$).

Sugars play two important roles in biology: they function (i) as an energy source (glucose, fructose) and (ii) as building material for the cell (glucose, ribose). What role the sugar plays depends on the type of bonding between the links in the sugar chain (glycosidic linkage); one linkage type generates the building material cellulose, the other the energy source glycogen, which is ultimately decomposed by the cell. Another class of polysaccharides is the riboses, which play an important part in the formation and stabilization of the DNA double helix.

Glycoproteins

Polysaccharides can connect with proteins to form glycoproteins. The vast majority of therapeutic proteins are glycoproteins (Table 1.1). The sugar structures attached to the proteins are manifold. They determine the functionality of the protein as the location and type of integration dominate folding and flexibility of the protein. Both the type and function of this glycosylation are species specific – a protein expressed in a yeast cell displays sugar structures different from those expressed in an animal cell. As this glycosylation pattern does not coincide with that of humans, those proteins usually show different biological activity and immunogenicity from those expressed in human cells.

Lipopolysaccharides

Molecular combinations of sugars with lipids (glycolipids) form a large part of the cell wall of Gram-negative bacteria (lipopolysaccharides).

1.3.2.4 **Lipids**

Lipids have a water-repelling as well as a water-attracting side, which makes them ideal structural elements of cytoplasmic membranes. Figure 1.17 shows the structure of such a cytoplasmic membrane and the role of lipids. The water-soluble parts are directed to the cytoplasm and the surrounding aqueous phase. The water-repellant parts are conjunct inside the membrane. Therefore, water-soluble substances cannot simply penetrate the cell wall. Proteins are channeled in a controlled fashion by transport structures embedded in the cell wall. Lipids are formed by fatty acids connected by alcohol molecules.

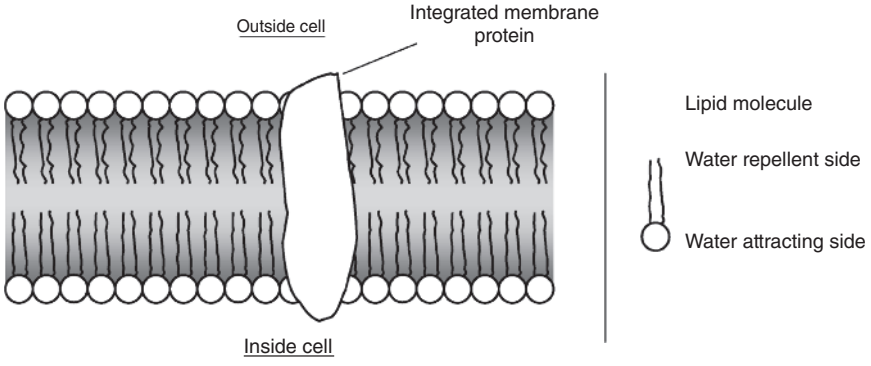

Figure 1.17 Schematic of a cytoplasmic membrane with lipids.