1

Biopharmaceuticals, an Introductory Overview

1.1 Introduction to Pharmaceutical Products

Pharmaceutical substances form the backbone of modern medicinal therapy. Such drugs may be categorized in a number of ways, such as by intended therapeutic application (e.g. anticancer agents, anticoagulants, or anesthetics), by their chemical structure, by their mode of action, or by how they are synthesized.

Most traditional pharmaceuticals are low molecular mass, organic chemicals (Table 1.1). Although some (e.g. aspirin) were originally isolated from biological sources, most are now manufactured by direct chemical synthesis, for various economic and/or technical reasons. Two types of manufacturing company thus comprise the "traditional" pharmaceutical sector: the chemical synthesis facilities, which manufacture the raw chemical ingredients ("fine chemicals") in bulk quantities; and the finished product pharmaceutical facilities, which purchase these raw bulk ingredients, formulate them into final pharmaceutical products, and supply these products to the market.

In addition to chemically synthesized drugs, a range of pharmaceutical substances are still obtained by direct extraction from biological material that produces these substances naturally. Such products, some major examples of which are listed in Table 1.2, may thus be described as products of traditional biotechnology. More recently, many pharmaceuticals produced in or by

Table 1.1 Some traditional pharmaceutical substances that are generally produced by direct chemical synthesis.

Drug	Molecular formula	Molecular mass (Da)	Therapeutic indication
Acetaminophen (paracetamol)	$C_8H_9NO_2$	151.16	Analgesic
Ketamine	$C_{13}H_{16}ClNO$	237.74	Anesthetic
Levamisole	$C_{11}H_{12}N_2S$	204.31	Anthelmintic
Diazoxide	$C_8H_7ClN_2O_2S$	230.7	Antihypertensive
Acyclovir	$C_8H_{11}N_5O_3$	225.2	Antiviral agent
Zidovudine	$C_{10}H_{13}N_5O_4$	267.2	Antiviral agent
Dexamethasone	$C_{22}H_{29}FO_5$	392.5	Anti-inflammatory and immunosuppressive agent
Misoprostol	$C_{22}H_{38}O_5$	382.5	Anti-ulcer agent
Cimetidine	$C_{10}H_{16}N_6$	252.3	Anti-ulcer agent

Biopharmaceuticals: Biochemistry and Biotechnology, Third Edition. Gary Walsh.

Table 1.2 Some pharmaceuticals that were traditionally obtained by direct extraction from biological source material. Many of the protein-based pharmaceuticals listed are now also produced by genetic engineering.

Substance	Medical application
Blood products (e.g. coagulation factors)	Treatment of blood disorders such as hemophilia A or B
Vaccines	Vaccination against various diseases
Antibodies	Passive immunization against various diseases
Insulin	Treatment of diabetes mellitus
Enzymes	Thrombolytic agents, digestive aids, debriding agents (i.e. cleansing of wounds)
Antibiotics	Treatment against various infectious agents
Plant extracts (e.g. alkaloids)	Various, including pain relief

engineered biological systems have come to market, and these have been called "biopharmaceuticals." The term "biopharmaceutical" was first used in the 1980s to describe therapeutic proteins produced by modern biotechnological techniques, specifically via genetic engineering (Chapter 5) or, in the case of monoclonal antibodies at that time, by hybridoma technology (Chapter 9).

Although the majority of biopharmaceuticals now approved or in development are proteins (Box 1.1 and Chapter 4) produced via genetic engineering, this term now also encompasses nucleic-acid-based, i.e. deoxyribonucleic acid (DNA)- or ribonucleic acid (RNA)-based products, as well as engineered whole-cell-based products (Chapter 16).

Box 1.1 Overview of protein structure

Proteins are large molecules made up of one or more chains of amino acids, known as polypeptides. These amino acids are linked by peptide bonds, and the sequence of these amino acids is determined by the gene encoding the polypeptide. Once synthesized, the polypeptide chain folds into a unique three-dimensional shape, which is crucial for its biological function. This specific shape is stabilized by various weak interactions, such as hydrogen bonds, and is dependent on the amino acid sequence. Any changes in the sequence can alter the shape and, consequently, the function of the protein.

Protein structure can be described at four levels:

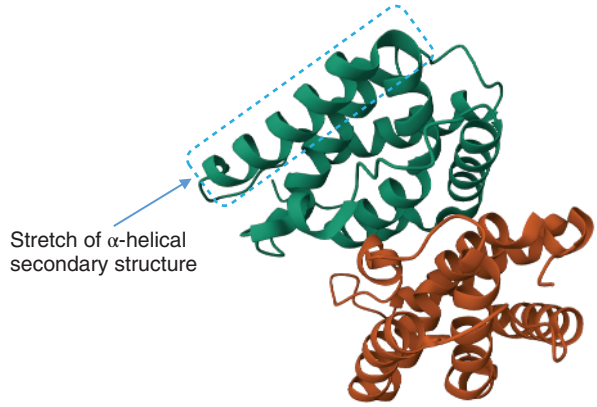
- **Primary structure:** The specific sequence of amino acids in the polypeptide chain, including the positioning of any disulfide bonds.
- **Secondary structure:** Regular patterns formed by adjacent amino acids, such as alpha-helices and beta-strands, often over short sequences.
- **Tertiary structure:** The overall three-dimensional arrangement of all atoms in the polypeptide, including regions of secondary structure and less ordered areas like loops.
- **Quaternary structure:** The overall spatial arrangement of multiple polypeptide subunits in a protein composed of more than one polypeptide chain.

Most proteins from eukaryotes undergo covalent modifications during or after their synthesis on ribosomes. These modifications, known as co-translational and post-translational

modifications (PTMs), include processes like glycosylation, where sugar chains are attached to specific points on the polypeptide backbone.

Amino acid sequence of human interleukin 22 (IL-22; 179 amino acids, each represented by its single letter designation), and its associated 3-dimensional structure (conformation). Data courtesy of the NCBI (www.ncbi.org); Sequence information: GenBank, AAH69308.1. Structural information: Protein Data Bank (PDB) ID 1M4R.

```
MAALQKSVSSFLMGTLATSCLLLLALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLAKEASLADNNT
DVRLIGEKLFHGVSMSERCYLMKQVLNFTLEEVLFPQSDRFQPYMQEVVPFLARLSNRLSTCHIEGDDLH
IQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI
```



Terms such as "biologic," "biopharmaceutical" and "products of pharmaceutical biotechnology," or "biotechnology medicines" have now become an accepted part of the pharmaceutical lexicon. These terms are sometimes used interchangeably but can mean different things to different people.

A summary of their meaning, as used in this book, is provided in Table 1.3. In reality, such designations can be somewhat artificial. For example, insulin products extracted directly from the pancreas of slaughtered pigs would be considered as products of traditional biotechnology, whereas insulin products produced via genetic engineering would be considered as biopharmaceuticals. In addition, a small number of RNA-based products (siRNA and antisense products – Chapter 16) are often described as biopharmaceuticals, even though they are manufactured by direct chemical synthesis.

The pace of new drug discovery has grown significantly over the past half-century in particular. Biomedical research continues to broaden our understanding of the molecular mechanisms underlining both health and disease, facilitating a knowledge-based approach to drug identification and discovery. Disciplines such as genomics, proteomics, and bioinformatics help identify both new potential drugs, and in particular new potential drug targets (Chapter 2). Advances in analytical (bio)chemistry have led to the development of a range of technologies and instruments (for example, mass spectrometry) capable of detecting, quantifying, and elucidating the structure of biomolecules associated with health and disease. Some such advances have facilitated the rapid analysis of multiple samples (high-throughput screening), speeding up the screening and identification of both new drugs and/or drug targets very significantly. A potent example of how such scientific and

Table 1.3 The meaning of the various indicated terms, as used in this book.

Term	Meaning
Biotechnology	The use of biological systems (for example, cells or tissues) or biological molecules (for example, enzymes or antibodies) for/in the development or manufacture of commercial products.
Genetic engineering	Knowledge-based laboratory processes which alter the DNA makeup of an organism (for example, the introduction of a gene coding for a therapeutic protein into a microbial or other cell, which can then synthesize the "recombinant" therapeutic protein in large amounts).
Pharmaceutical products of biotechnology (biotechnological products or biologics)	Any active pharmaceutical ingredient (API) manufactured via biotechnology. Broad term incorporating: • APIs extracted from biological material that produces them naturally (for example, insulin extracted from porcine pancreatic tissue); • APIs composed of or produced in engineered biological systems, including engineered cells, tissue, or whole animals/plants (for example, recombinant therapeutic proteins produced in genetically engineered cells or in transgenic plants or transgenic animals); • APIs produced *in vitro* in which the sole or a primary manufacturing step uses cells/tissues/viruses or components thereof such as enzymes (for example, the production of mRNA vaccines using RNA polymerase enzymes).
Pharmaceutical products of traditional biotechnology (traditional biotechnological products)	Any API produced in and extracted from any biological source material that produces that substance naturally (for example, antibiotics produced naturally by specific microorganisms, blood clotting factors extracted directly from donated blood or insulin preparations obtained by extraction for the pancreatic tissue of slaughter animals).
Biopharmaceuticals (pharmaceutical products of modern biotechnology)	Any API which is: • Itself a genetically engineered product (for example, genetically engineered cells or viruses used for gene therapy) or: • Produced using a genetically engineered biological system (for example, recombinant therapeutic proteins produced in genetically engineered cells) or: • Produced using engineered biological molecules (for example, mRNA vaccines produced via in vitro transcription, using a recombinant RNA polymerase enzyme).

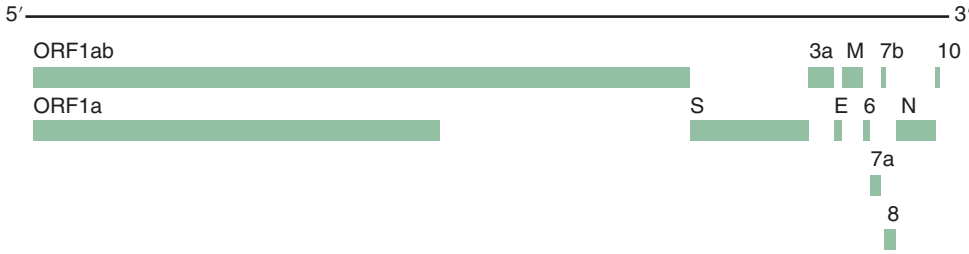
technological advances can lead to rapid drug development was most recently provided by the COVID-19 pandemic, with the development of several effective vaccines and therapies within months of causative agent identification (Box 1.2).

Box 1.2 The rapid development of COVID-19 vaccines and therapeutics

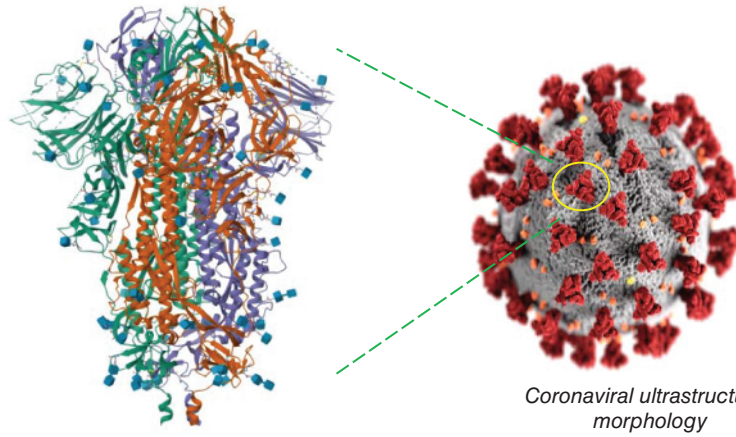
The rapid development of multiple vaccines and therapeutic agents against the COVID-19 causative agent (SARS-CoV-2) illustrates how advances in scientific techniques and knowledge can facilitate the rapid, knowledge-based identification of drug targets and associated drug development.

The first officially recorded cases of COVID-19 were reported in China in December 2019, and the SARS-CoV-2 coronavirus was identified as its causative agent in that same month.

The full genome sequence of the original Wuhan viral strain (Wuhan-Hu-1; a 29,903 base, single-stranded RNA virus) was published in GenBank (the US National Institutes of Health's genetic sequence database) on January 13, 2020 (https://www.ncbi.nlm.nih.gov/genbank/; accession sequence NC_045512). Sequence analysis identified the various proteins encoded by the viral genome (open reading frames; ORFs):



Comparison of these sequences with previously deposited coronavirus sequences allowed rapid identification of the likely function of each such protein. It also allowed their recombinant production which, in turn, allowed further characterization of the individual proteins. Chief among these was the 1273 amino acid "S protein" – the spike protein found on the viral surface (coded by the "S" ORF shown above). As the major viral surface feature, the S protein became the target of vaccine development within days of genome sequence publication. Recombinant S protein production also allowed determination of its three-dimensional structure, as deposited in the NIH's Protein Data Bank (www.rcsb.org/, ID number 6VXX) on February 25, 2020. The S protein functions as a homotrimer (3 individual S polypeptides interacting together), giving the protein its characteristic crown-like appearance. As known from previously studied coronaviruses, the S protein specifically recognizes angiotensin-converting enzyme 2 (ACE2), found on the surface of viral-sensitive cells, with viral docking via ACE-2 binding initiating subsequent cellular entry. This knowledge also suggested that mAbs capable of binding the Viral S protein could be useful in the treatment/amelioration of COVID-19, by blocking viral interaction with its human ACE2 cell surface receptor.



Three dimensional structure of the viral S protein trimer (individual S protein polypeptides in different colours)

Coronaviral ultrastructural morphology

Alissa Eckert, Dan Higgins / U.S. Centers for Disease Control and Prevention / Public Domain

(Continued)

Box 1.2 (Continued)

Such detailed knowledge of viral structure and of how it interacts with human cells allowed rapid and knowledge-based development of effective vaccines and therapeutic agents. This knowledge base also made rapid regulatory approval of those products possible. Within a year of initial viral characterization 2 COVID- vaccines had gained approval for human use (trade-names Comirnaty and Spikevax). By mid-2022, a total of 11 vaccines or mAb-based products targeting COVID-19 had been approved in the EU and/or the United States (table below). Moreover, the ability to rapidly sequence the genomes of emerging SARS-CoV-2 viral strains allows researchers to rapidly develop modified vaccines and mAbs effective in the vaccination against or treatment of such emerging strains.

Product	Description	Use	Approved*
Nuvaxovid/Novavax	SARS-CoV-2 recombinant (full-length) spike protein	Vaccine (COVID-19)	2022 (US) 2021 (EU)
Spikevax (previously COVID-19 Vaccine Moderna)	Single-stranded mRNA encoding the (full-length) viral spike (S) protein of SARS-CoV-2	Vaccine (COVID-19)	2021 (EU) 2020 (US)
Jcovden (COVID-19 vaccine Janssen)	Recombinant, replication-deficient adenovirus type 26 encoding the SARS-CoV-2 spike glycoprotein	Vaccine (COVID-19)	2021 (EU)
Vaxzevria	Recombinant, replication-deficient chimpanzee adenovirus that encodes the SARS-CoV-2 spike protein	Vaccine (COVID-19)	2021 (EU)
Comirnaty (Pfizer BioNTech COVID-19 vaccine)	mRNA Vaccine. Single-stranded, 5′-capped messenger RNA produced using a cell-free in vitro transcription from the corresponding DNA templates, encoding the viral spike (S) protein of SARS-CoV-2	Vaccine (COVID-19)	2020 (US) 2020 (EU)
Bebtelovimab	mAb (IgG) that binds to an epitope within the receptor-binding domain of the spike protein of SARS-CoV-2	COVID-19 treatment	2022 (US)

Product	Description	Use	Approved*
Evusheld	Combination of 2 mAbs, directed against two (non-overlapping) epitopes on the receptor-binding domain of (SARS-CoV-2) spike protein	COVID-19 prevention and treatment	2022 (EU) 2021 (US)
Bamlanivimab and eteseviman	Combination of 2 mAbs that bind to distinct but overlapping epitopes within the receptor-binding domain of the spike protein of SARS-CoV2	COVID-19 treatment	2021 (US; authorization paused in 2022 as product not sufficiently effective against Omicron viral variant)
Regkirona	mAb targeting the spike protein of SARS-CoV-2	COVID-19 treatment	2021 (EU)
Ronapreve (EU)/ Regen-cov (US)	Combination of two mAbs targeting distinct epitopes of SARS-CoV 2 spike protein	Prevention and treatment of COVID-19	2021 (EU) 2020 (US; authorization paused in 2022 as product not sufficiently effective against Omicron viral variant)
Xevudy (Sotrovimab in US)	mAb targeting the receptor-binding domain (RBD) of the Covid-19 spike (S) protein	COVID-19 treatment	2021 (EU) 2021 (US; authorization paused in 2022 as product not sufficiently effective against Omicron viral variant)

*US approval under Emergency Use Authorisation (EUA), EU approval under Conditional Marketing Authorization; Chapter 2).
Illustrations used courtesy of NIH (www.ncbi.nlm.nih.gov) and CDC (https://phil.cdc.gov/, ID 23312).

The past several decades have also witnessed technological advances that facilitate the large-scale manufacture of pharmaceutical products. This is particularly true in the case of therapeutic proteins, where recombinant DNA technology (genetic engineering) can be used to produce such proteins in very large quantities (Figure 1.1).

1.2 Genetic Engineering and the Advent of Biopharmaceuticals

Research dating back to the early/mid-twentieth century began to identify a host of proteins produced naturally in the body with obvious therapeutic potential. Examples include interferons and

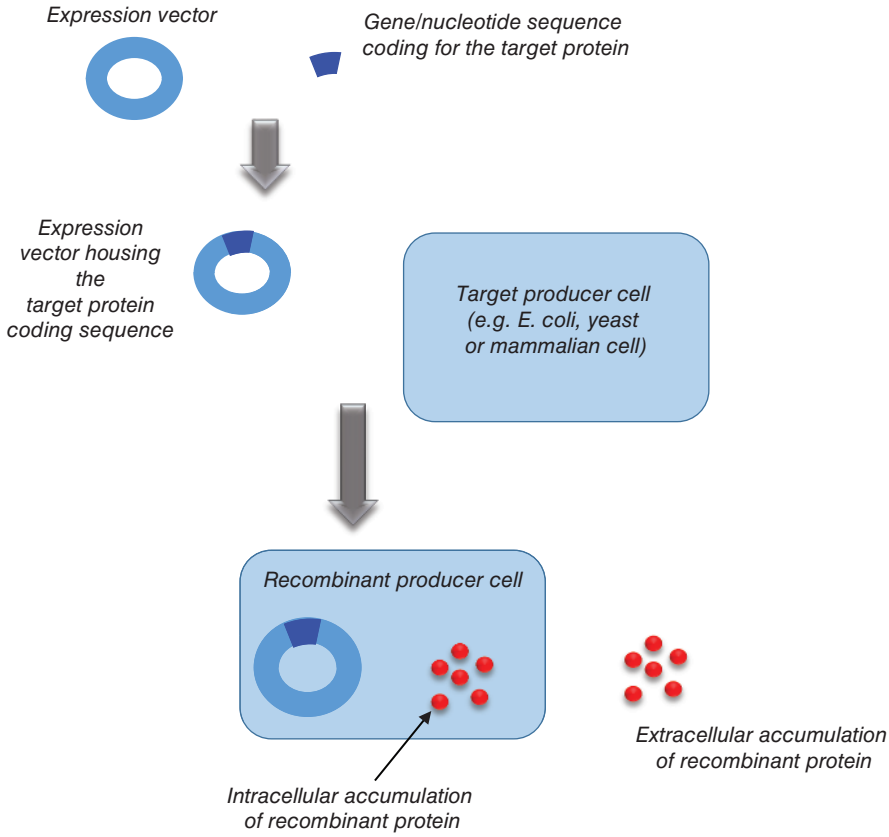


Figure 1.1 Overview of protein production via recombinant DNA technology (genetic engineering). The genetic sequence coding for the desired protein is introduced into an "expression vector" (usually a circular piece of DNA called a plasmid), which is then in turn introduced into a producer cell (often *E. coli*, a yeast, or a mammalian cell such as Chinese hamster ovary (CHO) cells). The now-engineered (recombinant) producer cells are grown (cultured) in large numbers. Genetic elements found in the expression vector drive high-level transcription and translation of the desired target protein's genetic sequence, i.e. drive high-level recombinant protein production. Depending on the producer system used, the resultant recombinant protein may accumulate within the cell (intracellular production), or may be exported from the cell into the extracellular environment. Extracellular production is usually preferred, as it allows subsequent recovery of the desired protein without a need to collect and break open (lyse) the producer cells. Recombinant protein production is discussed in detail in Chapter 5.

interleukins (which regulate, in particular, the immune response) and growth factors such as erythropoietin (EPO; which stimulates red blood cell production).

The widespread medical application of many such proteins was rendered impractical due to the tiny quantities in which they are naturally produced. The advent (in the 1970s) of recombinant DNA technology (genetic engineering; Chapter 5) and monoclonal antibody technology (hybridoma technology; Chapter 9) overcame many such difficulties, and marked the birth of the biopharmaceutical sector.

Recombinant DNA technology has had a fourfold positive impact upon the production of pharmaceutically important proteins:

- *It overcomes the problem of source availability*. Many proteins of therapeutic potential are produced naturally in the body in minute quantities. Examples include interferons (Chapter 11), interleukins (Chapter 12), and colony-stimulating factors (CSFs; Chapter 13). This rendered impractical their direct extraction from native source material in quantities sufficient to meet likely clinical demand. Recombinant production (Chapter 5) allows the manufacture of virtually any protein in large quantities.
- *It overcomes problems of product safety*. Direct extraction of product from some native biological sources has, in the past, led to the unwitting transmission of disease. Examples include the transmission of blood-borne pathogens such as hepatitis B and C and human immunodeficiency virus (HIV) via infected blood products and the transmission of Creutzfeldt–Jakob disease to persons receiving growth hormone (GH) preparations derived from human pituitaries.
- *It provides an alternative to direct extraction from inappropriate/dangerous source material*. A number of therapeutic proteins were traditionally extracted from human urine. Follicle-stimulating hormone (FSH), the fertility hormone, for example, was obtained from the urine of postmenopausal women, and a related hormone, human chorionic gonadotrophin (hCG), was extracted from the urine of pregnant women (Chapter 14). Urine is not considered a particularly desirable source of pharmaceutical products. While a few products obtained from this source remain on the market in some world regions, most such products are now produced by recombinant means. Other potential biopharmaceuticals are produced naturally in downright dangerous sources. Ancrod, for example, is a protein displaying anticoagulant activity and, hence, is of potential clinical use. It is, however, produced naturally by the Malaysian pit viper. Although retrieval by milking snake venom is possible, and indeed may be quite an exciting procedure, recombinant production in less dangerous organisms, such as *Escherichia coli* or *Saccharomyces cerevisiae*, would be considered preferable.
- *It facilitates the generation of engineered therapeutic proteins displaying some clinical or other advantage over the native protein product*. Techniques such as site-directed mutagenesis facilitate knowledge-based introduction of predefined changes in a protein's amino acid sequence. Such changes can be as minimal as the insertion, deletion, or alteration of a single amino acid residue, or can be more substantial (e.g. the alteration/deletion of an entire domain, or the generation of a novel hybrid protein). Such changes can be made for a number of reasons, and the majority of protein-based biopharmaceuticals now on the market are engineered in some way. An overview of selected protein engineered products is provided in Table 1.4. It is also worth noting that some biopharmaceuticals are engineered by means other than altering their amino acid sequence (Table 1.5). These and other examples will be discussed in subsequent chapters.

Despite the undoubted advantages of recombinant production, some protein-based products extracted directly from native source material remain on the market. In certain circumstances, direct extraction of native source material can prove equally or even more attractive than recombinant production. This may be for an economic reason if, for example, the protein is produced in very large quantities by the native source and is easy to extract/purify, e.g. human serum albumin (HSA). Also,

Table 1.4 Selected protein-engineered biopharmaceutical types/products that have now gained marketing approval. These and additional such products will be discussed in detail in subsequent chapters.

Product description/type	Alteration introduced	Rationale
Faster-acting insulins (Chapter 14)	Modified amino acid sequence	Generation of faster-acting insulin
Slow-acting insulins (Chapter 14)	Modified amino acid sequence	Generation of slow-acting insulin
Modified tissue plasminogen activator (tPA; Chapter 15)	Removal of three of the five native domains of tPA	Generation of a faster-acting thrombolytic (clot degrading) agent
Modified blood factor VIII (Chapter 15)	Deletion of one domain of native factor VIII	Production of a lower molecular mass product
Chimeric/humanized antibodies (Chapter 19)	Replacement of most/virtually all of the murine amino acid sequences with sequences found in human antibodies	Greatly reduced/eliminated immunogenicity. Ability to activate human effector functions
"Ontak," a fusion protein (Chapter 9)	Fusion protein consisting of the diphtheria toxin linked to interleukin-2 (IL-2)	Targets toxin selectively to cells expressing an IL-2 receptor

Table 1.5 Selected examples of protein biopharmaceuticals engineered by means other than alteration of their amino acid sequence. These and/or additional examples will be discussed in more detail in subsequent chapters.

Product	Alteration introduced	Rationale
Esperoct (recombinant coagulation factor VIII, used to treat hemophilia A)	Covalent attachment of polyethylene glycol (PEG)	The PEG moiety extends the plasma half-life of the product
Sogroya (recombinant human growth hormone used to treat growth hormone deficiency)	Covalent attachment of a fatty acid to the protein backbone	The fatty acid moiety extends the plasma half-life of the product
Padcev (antibody–drug conjugate (ADC) used to treat urothelial cancer)	Attachment of a toxin (monomethyl auristatin E; MMAE) to the protein backbone	Toxin kills the cancer cells targeted by the antibody
Rybrevant (a monoclonal antibody used to treat advanced non-small-cell lung cancer)	Alteration of the mAb's carbohydrate side chain such that it contains only low levels of the sugar fucose.	Reduction in the fucose levels found in a mab's sugar side chain is associated with increased antibody-dependent cellular cytotoxicity (ADCC), and hence more effective destruction of the target cancer cells

some blood factor preparations purified from donor blood actually contain several different blood factors and, hence, can be used to treat several hemophilia patient types. Recombinant blood factor preparations, on the other hand, contain but a single blood factor and, hence, can generally be used to treat only one hemophilia type (Chapter 15). The most lucrative such protein still produced via purification from its naturally producing source is Botulinum toxin type A (Botox; Box 1.3).

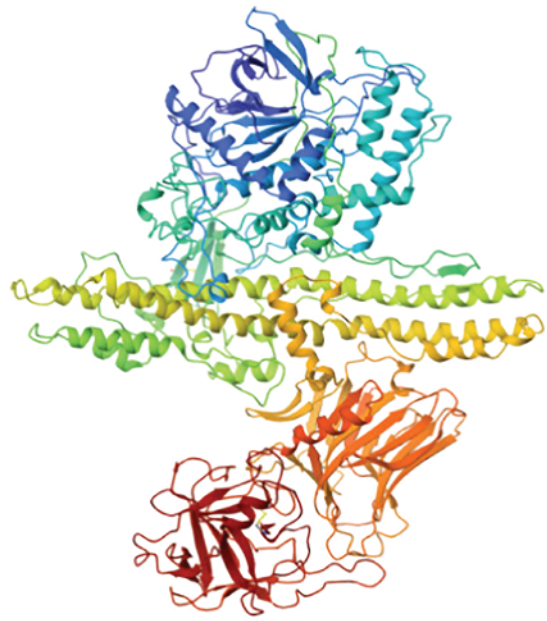
Box 1.3 Botox

Produced by the anaerobic bacterium *Clostridium botulinum*, botulinum toxin is the most potent microbial toxin known. Several toxin variants exist (names A to G), with type A being the most studied and medically used. Type A toxin has an estimated lethal dose in humans of 1.5 – 2.0 ng/kg body weight if administered intravenously. It is a neurotoxin, bringing about its effects by blocking the release of the neurotransmitter acetylcholine at neuromuscular junctions, thereby causing flaccid paralysis in which the muscle remains in the relaxed (limp) state.

Type A toxin is produced as a 150 kDa, 1295 amino acid polypeptide chain (diagram below). Upon bacterial release it is proteolytically cleaved into a 2-fragment protein, of approximately 100 kDa (the heavy chain) and 50 kDa (the light chain), respectively, held together by a disulfide bond. The heavy chain contains a binding domain, specific for receptor molecules (gangliosides) present in the plasma membrane of target neuronal cells. Binding is followed by toxin internalization into the neuronal cell. The light chain displays proteolytic activity against neuronal cytoplasmic proteins involved in promoting acetylcholine release. Blockage of acetylcholine release into the neuromuscular junction thereby prevents neurotransmission from the neuron to the muscle.

Type A botulinum toxin has found numerous medical and cosmetic applications, with the underlining mode of action being to relax muscles. Medically, it is used, for example, to treat spasticity associated with strokes or spinal cord injury, overactive bladder, and cervical dystonia. Cosmetically it is used to reduce facial wrinkles. Commercially, it is sold under various tradenames, including Botox and Dysport.

Approved for medical use since 1989, Botox is produced by fermentation of the Hall strain of *Clostridium botulinum*, which produces the type A toxin naturally. The toxin is then purified from the culture solution by dialysis and several acid precipitation steps, before addition of human albumin as an excipient. After sterile filtration through a 0.2 μm filter, it is filled into pre-sterile vials and vacuum dried to produce the final product. By the mid-2020s Botox was generating annual global revenues of some US$ 5 billion, making it one of the top-selling products of biotechnology.



Three dimensional structure of botulinum neurotoxin serotype A (PDB entry 3BTA)

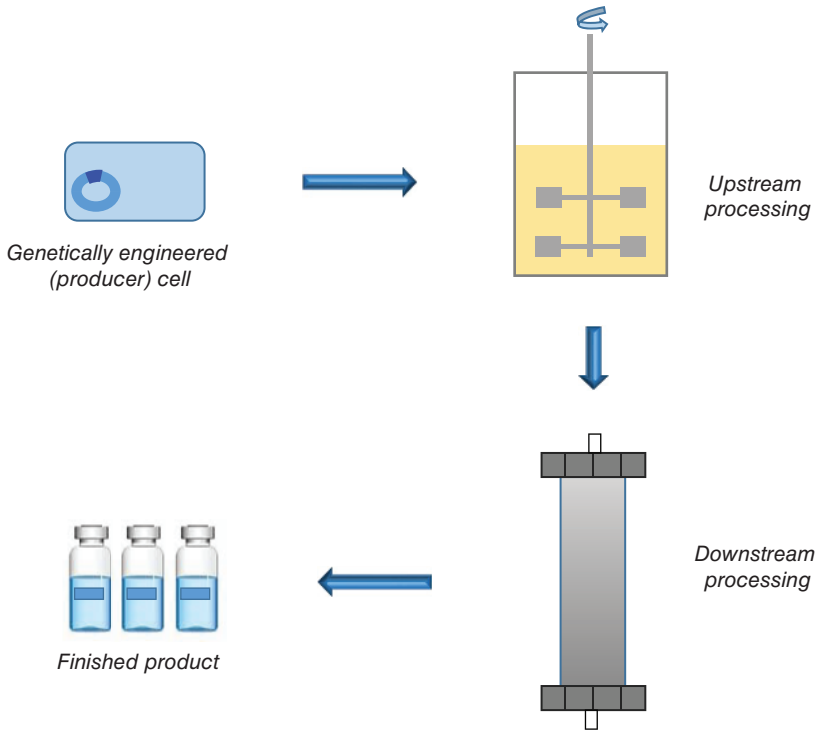


Figure 1.2 Recombinant therapeutic protein manufacture in overview. Refer to text for details.

1.2.1 Manufacture of Recombinant Therapeutic Proteins

As discussed later in this chapter, some 95% of all biopharmaceuticals currently on the market are proteins produced via recombinant DNA technology. During initial product development, the gene/nucleotide sequence coding for the therapeutic protein of interest will be introduced into a recombinant expression system (genetically engineered cell line), as overviewed in Figure 1.1 and considered in detail in Chapter 5. Common expression systems used include bacterial cells (mainly *E. coli*), yeast (mainly *Saccharomyces cerevisiae* or *Pichia pastoris*), and mammalian cells (mainly Chinese Hamster Ovary (CHO) cells). The exact cell type used will depend upon a number of factors, such as if the therapeutic protein requires a post translational modification, such as glycosylation (Chapter 5).

Once approved for general medical use (Chapter 2), routine product manufacture (Figure 1.2 and Chapters 6 and 7) entails an upstream processing phase (in which the genetically engineered producer cells are grown on an industrial scale), followed by a downstream processing phase (incorporating product extraction/recovery, purification, formulation, and fill into final product containers).

1.3 Biopharmaceuticals: Current Status and Future Prospects

The first biopharmaceutical to come on the market was a recombinant human insulin produced in *E. coli* (tradename humulin). It was initially approved in the United States by the FDA in 1982.

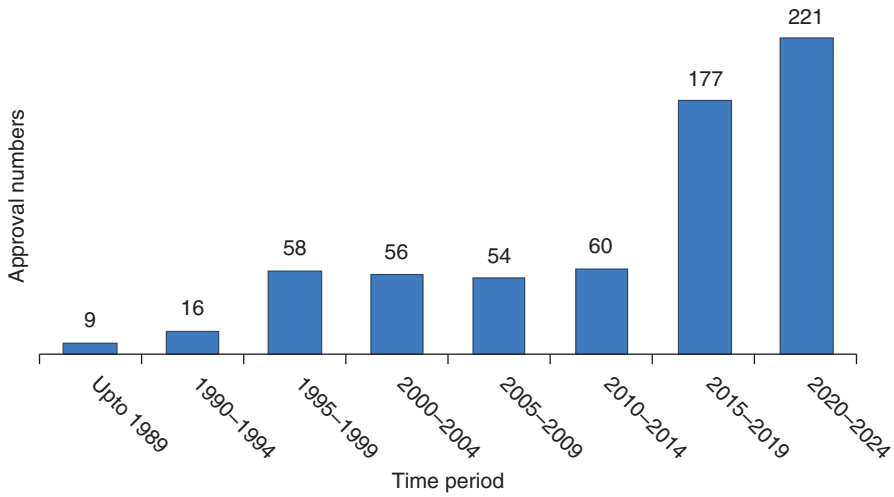


Figure 1.3 Number of biopharmaceutical products approved in the United States and/or the EU during the indicated time periods. Walsh, G. and Walsh, E. (2022) / Springer Nature.

The number of approvals has grown steadily over the intervening years (Figure 1.3) and by the mid-2020s some 650 such products had come on the market within the United States and/or the European Union (EU). Typically, in the region of 30% of all new drugs coming on the market in any given year are biopharmaceuticals. In absolute numbers this usually equates to between 20 and 40 individual products entering the market annually.

Approved products include a range of hormones, blood factors, and thrombolytic agents, as well as monoclonal antibodies, vaccines, nucleic acids, and cell-based products (Figure 1.4). The most common indications (approved therapeutic use) for biopharmaceuticals include cancer, inflammatory-related conditions, and vaccination against infectious diseases.

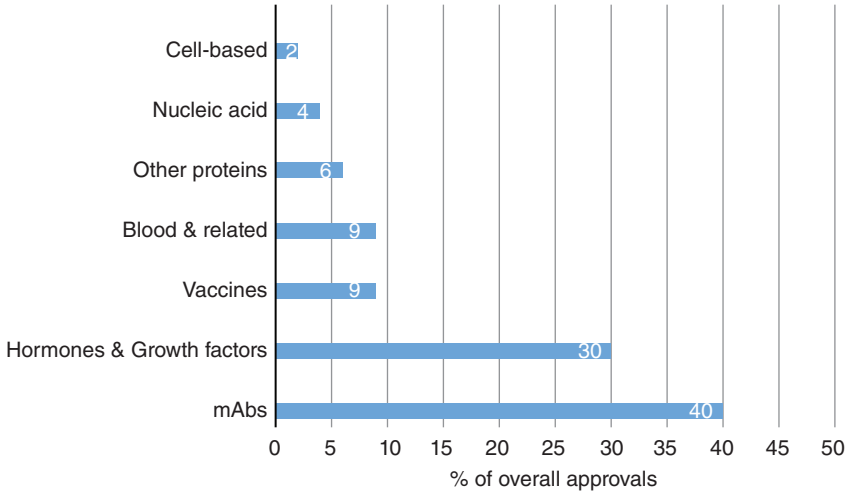


Figure 1.4 Overall biopharmaceuticals approvals profile by product category.

1.3.1 Monoclonal Antibodies

In terms of product numbers, monoclonal antibodies (mAbs) represent the most significant biopharmaceutical product category (Chapter 9). In any given year they generally constitute just over half of all product approvals. This dominance stems from the fact that such antibodies can be produced to selectively bind to almost any target molecule (antigen), and thus they can find application in the treatment of broad range of diseases or medical conditions. In overview, the majority of mAb-based products bring about their therapeutic effects either by binding to, and thereby inactivating, a target substance whose over-production in the body is causing or acerbating a disease, or by binding to specific cell types (e.g. cancer cells), thereby triggering their destruction by additional elements of the immune system (Figure 1.5).

As a class of product, mAbs have also been subjected to a whole range of engineering approaches, which has further broadened their therapeutic use (Figure 1.6).

1.3.2 Nucleic Acid and Engineered Cell-Based Products

A small but increasing proportion of biopharmaceuticals approved to date are nucleic acid or engineered cell based (Chapter 16). Nucleic acid products can be subcategorized into RNA-based products and gene therapy products, as summarized in Figure 1.7.

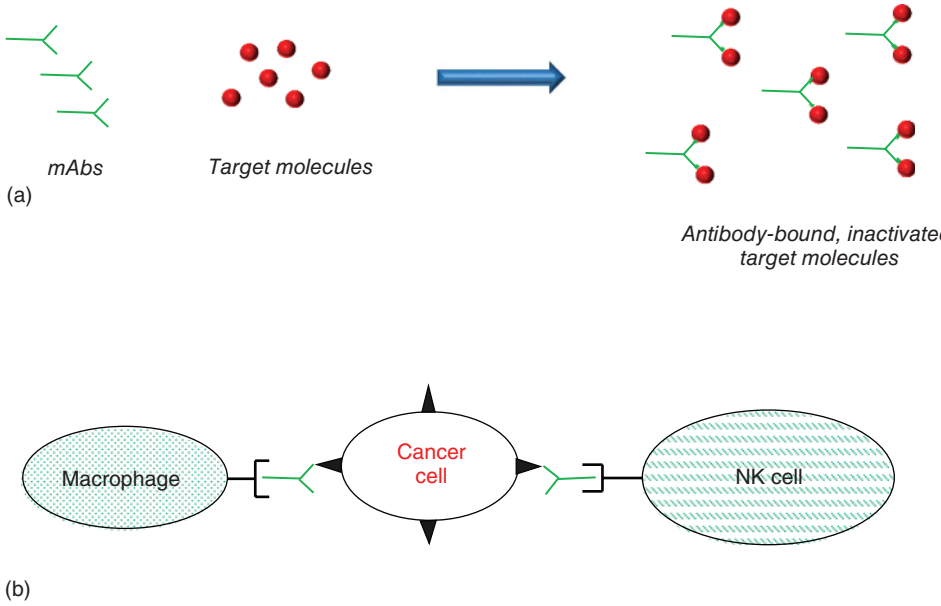


Figure 1.5 An overview of the mechanism of action of most therapeutic antibodies. Some medical conditions are caused by over-production of signaling molecules in the body (for example, overproduction of the pro-inflammatory protein TNF can trigger various inflammatory conditions such as rheumatoid arthritis). Administration of an antibody capable of binding to such target molecules can treat/ameliorate the condition as antibody binding will inactivate them, (a). Many anti-cancer antibodies bring about their therapeutic effect by selectively binding to the surface of cancer cells. This in turn facilitates additional immune system cells such as macrophages and natural killer (NK) cells to dock at the surface of the cancer cell (by binding to a different part of the antibody). The docked immune cells will then trigger the destruction of the cancer cell via various immunological mechanisms, (b).

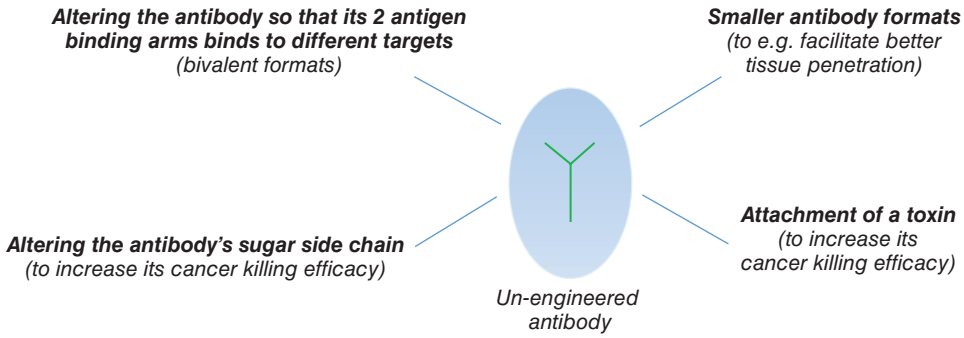


Figure 1.6 An overview of various engineering approaches adopted in mAb technology. Refer to Chapter 9 for further detail.

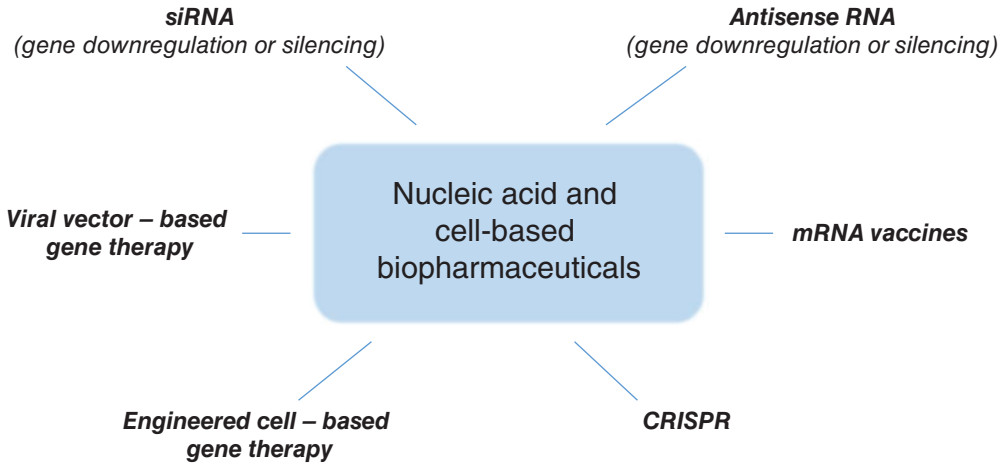


Figure 1.7 Overview summary of the various categories of nucleic acid and engineered cell–based pharmaceuticals thus far approved for medical use. Refer to text and Chapters 10 and 16 for further detail.

Many disease states are associated with the inappropriate production or overproduction of specific gene products. One approach to treating such conditions could entail the reduction (downregulation) or cessation (silencing) of such gene expression. This could be achieved by blocking target gene transcription or translation. Two nucleic-acid-based product types aim to achieve just that – antisense RNA molecules and small interfering ribonucleic acid (siRNA) molecules, with some 10 antisense and 5 siRNA products having gained approval for medical use in the EU and/or United States by the early 2020s. Both product types are RNA-based molecules of a nucleotide sequence specifically designed to form a complementary duplex with target RNA molecules (via Watson–Crick complementary base pairing). This in turn triggers cellular destruction of the target RNA, and hence gene downregulation or silencing. Notably, these RNA products are manufactured by direct chemical synthesis, and as such do not strictly fall into the description of a biopharmaceutical, although they have historically been described as such. Their direct chemical synthesis does, however, facilitate the chemical modification of their nucleotide backbone, which makes them more resistant to nuclease enzymes and hence resistant to immediate degradation by endogenous nucleases upon their administration.

mRNA-based vaccines, such as the COVID-19 vaccine products Spikevax and Comirnaty, constitute another group of RNA-based products. Moreover, these products can be described as

biopharmaceuticals (Table 1.3) as they are synthesized via an in vitro transcription process utilizing the enzyme DNA polymerase (Chapter 10). mRNA vaccines invariably consist of a mRNA molecule which encodes part or all of a prominent viral (or other pathogen) surface antigen, such as the spike (S) protein of the COVID-19 virus. Intramuscular administration results in local cellular uptake and subsequent translation of the mRNA into surface antigen protein, which in turn triggers an immune response, and hence vaccination.

Gene therapy, as the name suggests, entails the introduction of a gene into the genetic complement of a cell, such that subsequent *in vivo* expression of the gene achieves a therapeutic goal (Chapter 16). Gene therapy is predominantly used to:

- Treat genetic diseases: Such conditions are invariably caused by the sufferer inheriting mutated gene(s) whose gene product (usually a protein) is either not produced in sufficient quantities or, (because of an altered amino acid sequence), is poorly functional/non-functional. Gene therapy facilitates the introduction and expression of a copy of the gene(s) coding for fully functional copies of the protein in question.
- Introduce a new gene into body cells whose expression treats a target condition such as cancer or an infectious disease. An example is CAR-T cell technology, as overviewed later in this section and as discussed in detail in Chapter 16.

While conceptually straightforward, the successful application of gene therapy has been retarded by multiple technical hurdles, including how to successfully deliver the therapeutic gene to the target cell population and how to subsequently achieve long-term gene expression (Chapter 16). Scientific developments continue to overcome many such challenges, at least in part, with 20 such products being approved for medical use in the EU and/or United States by the mid-2020s.

Approved products largely rely on one of 2 mechanisms to facilitate delivery of the therapeutic gene into the body (Figure 1.8);

- the packaging of the gene into an (non-replicating, non-pathogenic) engineered virus, which delivers the gene to body cells upon administration to the patient.
- the removal of target cells from the body, followed by their in vitro engineering (introduction of the therapeutic gene), followed by their reintroduction into the body.

A particularly prominent form of gene therapy is that of CAR-T cell immunotherapy used to treat various cancers. This entails the initial isolation of T-lymphocytes from the blood of the cancer sufferer. A synthetic gene coding for a "chimeric antigen receptor" (CAR) protein is introduced into the T-cells in vitro, followed by re-introduction of the genetically engineered T cells back into the patient's own body. The CAR protein, when expressed on the surface of the engineered T cells, is capable of binding to cancer cells and triggering their T-cell-mediated destruction (Chapter 16).

1.3.3 Biosimilars

Once a biopharmaceutical product loses patent protection it is open to other companies to develop and market copies of that product. Biopharmaceuticals are synthesized in biological systems and are structurally complex (Table 1.6). This makes it all but impossible to generate a product guaranteed to be identical to the original product. In most instances, however, it is possible to develop a copy that is structurally very substantially similar to the original ("reference") biopharmaceutical,

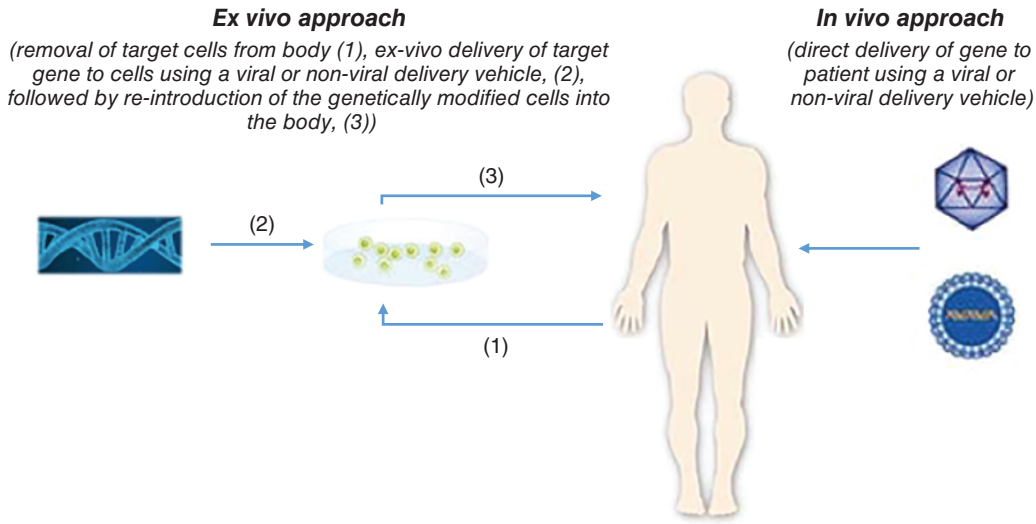


Figure 1.8 The main approaches by which gene therapy may be achieved in practice. Adapted from www.fda.gov.

Table 1.6 Some comparative properties of traditional "fine-chemical"-based pharmaceutical products versus biopharmaceuticals in relation to structure, manufacture, and analysis.

Property	Traditional fine chemical	Biopharmaceutical
Mass	Usually 0.1–5.0 kDa	5–150 kDa
Structure	Precisely known	Complex and can be heterogeneous if carrying a post-translational modification
Synthesis	Chemical, often straightforward	Biological, complex
Purification	Often straightforward	Usually complex
Analysis	Often straightforward	Usually complex
Immunogenicity	Often low	Potentially high

such that no clinically meaningful differences distinguish them. Such copies are termed "biosimilars," and their development and regulation is considered in Chapter 2.

Virtually all biopharmaceutical products that have lost patent protection to date are protein based. Therefore, it is unsurprising that biosimilars thus far approved are proteins. The first such products approved were copies of recombinant human growth hormone (hGH), approved initially in the EU in 2006, under the tradenames Omnitrope and Valtropin. Since then a variety of biosimilar EPOs, granulocyte colony-stimulating factors (G-CSFs), insulin, FSH, and mAbs have come on the market, with over 100 such products having been approved for use in the EU and/or United States by the mid-2020s (Table 1.7).

As the cost of biosimilar drug development is significantly lower than that for genuinely new biopharmaceuticals (Chapter 2), such products are generally priced at a 20–50% discount relative

Table 1.7 Selected examples of biosimilar products now approved for medical use in the EU and/or the United States. These and/or additional examples will be discussed in more detail in later chapters.

Biosimilar (tradename)	Description	Reference product (tradename)	Year first approved for medical use
Omnitrope	Human growth hormone (hGH), used to treat growth disturbances	Genotropin	2006
Retacrit	Erythropoietin (EPO), used to treat anemia	Eprex/Erypo	2007
Releuko	Granulocyte colony-stimulating factor (G-CSF) used to treat neutropenia	Neupogen	2022
Rezvoglar	Long-acting engineered insulin used to treat diabetes	Lantus	2021
Avsola	Monoclonal antibody, used to treat arthritis and some other inflammatory conditions	Remicade	2019
Kauliv	Analogue of human parathyroid hormone, used to treat osteoporosis	Forsteo	2023
Eydenzelt	Vascular endothelial growth factor (VEGF)-binding Fc fusion protein. Used to treat certain eye conditions, most notably neovascular (wet) age-related macular degeneration (AMD)	Eylea	2025

to the originator biopharmaceutical product. As biopharmaceuticals are among the most expensive of all drugs, savings can be considerable and biosimilar products command a relatively modest but rapidly increasing market share of the biopharmaceutical sector.

1.3.4 Market Value

Total annual global sales for biopharmaceuticals surpassed US$ 450 billion by the mid-2020s. This represents some 27% of the total global pharmaceutical sales at that time, while 7 of the top 10 pharmaceutical products sold (by product revenue) were biopharmaceuticals. mAbs represented the single most lucrative category of biopharmaceutical, generating almost two-thirds of all biopharmaceutical revenue. From a commercial perspective, revenues generated by biosimilar nucleic acid and engineered cell-based products remained relatively modest, generating collective global sales in the region of $20 billion in the mid-2020s.

The top 10 selling biopharmaceuticals collectively generated earnings of $131 billion at that time (Table 1.8). Eight of those 10 products are mAb based.

As a product class, biopharmaceuticals remain among the most expensive of all pharmaceutical substances. The wholesale price of different biopharmaceuticals can vary widely, and the true cost of manufacture is often estimated to be a small fraction of sale price. The wholesale price of any product, however, reflects not only cost of manufacture, but also initial product development and approval costs, ongoing company overheads, market size, and investor or shareholder return.

Insulin, relatively speaking, is an inexpensive biopharmaceutical. This likely reflects its large market size, its relatively low manufacturing costs (achieved via high-level production in microbial systems), and relatively low initial development costs (given its long history of use). Although difficult to generalize as different diabetics are treated with various different insulin preparations, and in different quantities, the typical cost of one year's supply of insulin in the United States ranges from approximately $1,000–5,000. The direct manufacturing cost is likely in the $100–500 range.

Table 1.8 The 10 top selling biopharmaceutical products in the mid-2020s.

Ranking	Product	Description and therapeutic use	Sales value (US$ billions)	Company
1	**Keytruda**	Monoclonal antibody used to treat a number of cancer types	29.5	Merck
2	**Ozempic**	Glucagon-like peptide 1 (GLP-1) receptor agonist used to treat type 2 diabetes	17.5	Novo Nordisk
3	**Dupixent**	Monoclonal antibody used to treat a number of conditions including atopic dermatitis and asthma	14.0	Sanofi/ Regeneron
4	**Skyrizi**	Monoclonal antibody used to treat a number of inflammatory conditions including psoriasis and Crohn's disease	11.7	Abbvie
5	**Darzalex**	Monoclonal antibody used to treat multiple myeloma and amyloidosis	11.6	Johnson & Johnson
6	**Stelara**	Monoclonal antibody used to treat a number of inflammatory conditions, namely psoriasis, psoriatic arthritis, Crohn's disease, and ulcerative colitis	10.3	Johnson & Johnson
7	**Eylea**	Monoclonal antibody-based fusion protein, used to treat certain eye conditions, mainly age-related macular degeneration	9.5	Regeneron, Bayer
8	**Opdivo**	Monoclonal antibody used to treat a number of cancer types	9.3	Bristol-Myers Squibb
9	**Humira**	Monoclonal antibody used to treat various inflammatory conditions, including rheumatoid arthritis and psoriasis	9.0	AbbVie
10	**Gardasil/ Gardasil-9**	Recombinant subunit vaccine against diseases caused by human papillomavirus (HPV)	8.6	Merck

The list price of biopharmaceuticals tends to be considerably higher in the United States than many other world regions, where one year's supply of the same or equivalent insulin preparations can be under $1,000. In many world regions the actual cost to the patient is significantly below the list price, due to governmental health schemes or cost defrayment via health insurance schemes.

Manufacturing cost of monoclonal antibodies, relative to insulin, is higher. mAbs are much larger and more complex than insulin (molecular mass of 150 kDa versus 5.8 kDa), and are produced in mammalian as opposed to less costly microbial systems. In the United States, the average annualized cost of treatment with a mAb is just under $100,000, although this cost varies significantly for different mAbs products (ranging from some $ 2,400 to $ 800,000 in one case). In terms of price per gram of mAb active substance, this equates to a cost ranging from $5,000 to in excess of $114,000 in one case. The direct cost of manufacture of most such mAbs is likely to be broadly similar. The broad price variation is likely due to issues such as market size (patient numbers) and differing therapeutic regimes (the amount of mAb administered per dose and the total duration of treatment). In effect, the pharmaceutical company must recoup its fixed overhead costs, product development costs, direct manufacturing costs, and profit via the gross amount of product it sells.

This last point is also illustrated by gene therapy–based products. The cost of developing and manufacturing such products is likely to be somewhat higher compared to that of protein-based

products. Most gene therapy products approved to date are, however, indicated for rare genetic conditions or are used to treat more common conditions in very limited or exceptional circumstances. For such products, therefore, recouping costs and generating profit rely on sales to a small patient base. For example, the gene therapy product Zolgensma is used to treat spinal muscular atrophy, a rare genetic neuromuscular disorder with an incidence of approximately 1 in 10,000 live births in the United States. The once-off genetic treatment comes at a cost of €2.1 million per patient in the United States.

1.3.5 Future Prospects

The biopharmaceutical pipeline remains robust. An estimated 7,800 products are currently in clinical development, although this includes biosimilar products, as well as products already approved to treat a specific target condition now being evaluated to treat a different condition. The single largest sub-category of products in development is monoclonal antibodies, and the most common target indications are cancer, genetic disorders, cardiovascular diseases, and infectious diseases – all leading causes of mortality and/or morbidity, especially in the west. Several hundred nucleic acid and engineered cell-based products are also currently in clinical trials, so the proportion of these product categories entering the clinic is likely to increase with time. Biosimilar development also remains vibrant, spurred on by the increasing numbers of blockbuster drugs losing patent protection, the potential savings to healthcare systems, and the fact that lower-cost biopharmaceuticals would also increase patient accessibility, particularly in less affluent world regions. The impact of biopharmaceuticals on human healthcare is therefore likely to increase still further with time.

Sources of Additional Information

Most individual biopharmaceutical products are now supported by dedicated websites, usually bearing the product's tradename. For example, the dedicated website for the therapeutic mAb herceptin is: www.herceptin.com. Such product websites usually contain comprehensive information regarding the product in question. They are maintained by the company making or marketing the product, and therefore will contain authoritative and reliable information.

Biopharmaceuticals are assessed and/or approved for general medical use by various global regulatory agencies such as the Food and Drug Administration in the United States (www.fda.gov) and the European Medicines Agency (www.ema.europa.eu). Such agency websites will contain authoritative and reliable information, and are generally searchable by key word.

Numerous companies and organizations gather and publish financial and technical reports relating to the biopharma sector. These include, but are not limited to, La Marie publishing (https://lamerie.com/), Fierce pharma (www.fiercepharma.com), and statista (www.statista.com).

Biotech-facing industry organizations generally publish biopharma-focused articles and reports. Such organizations include The Pharmaceutical Research and Manufacturers of America (www.phrma.org) and the Biotechnology Innovation Organization (www.bio.org).

Some Journal Articles

Guan, X., *et al*. (2023). Advances in SARS-CoV-2 receptor-binding domain-based COVID-19 vaccines. *Expert Review of Vaccines*, 22 (1), 422–439.

Joshi, D. (2022). Biosimilars in oncology: Latest trends and regulatory status. *Pharmaceutics*, 14, 2721.

Monga, A. *et al*. (2025). Biosimilars: A critical review of development, regulatory landscape, and clinical implications. *AAPS PharmSciTech*, 26, 46 (https://doi.org/10.1208/s12249-025-03038-2).

O'Flaherty, R. *et al*. (2020). Mammalian cell culture for production of recombinant proteins: A review of the critical steps in their biomanufacturing. *Biotechnology Advances*, 43, 107552.

Ranbhor, R. (2025). Advancing monoclonal antibody manufacturing: Process optimization, cost reduction strategies, and emerging technologies. *Biologics: Targets and Therapy*, 19, 177–187.

Sayed, N. *et al*. (2022). Gene therapy: Comprehensive overview and therapeutic applications. *Life Sciences*, 294, 120375.

Walsh, G. and Walsh, E. (2022). Biopharmaceutical Benchmarks 2022. *Nature Biotechnology*, 40 (12), 1722–1760.

Wang, Z. *et al*. (2022). Development of therapeutic antibodies for the treatment of diseases. *Molecular Biomedicine*, 3, 35.