1
History of Industrial Biotechnology

Arnold L. Demain, Erick J. Vandamme, John Collins, and Klaus Buchholz

1.1
The Beginning of Industrial Microbiology

Microbes have been extremely important for life on Earth. They are the progenitors of all life on Earth and are the preeminent system to study evolution. They provide rapid generation times, genetic flexibility, unequaled experimental scale, and manageable study systems. Estimates indicate 5×10^{31} microbial cells exist with a weight of 50 quadrillion metric tons. More photosynthesis is accomplished by microbes than by green plants. More than 60% of the earth's biomass is that of microbes. Over 90% of the cells in human bodies are microorganisms. Sterile animals are less healthy than those colonized by microbes.

Long before their discovery, microorganisms were exploited to serve the needs and desires of humans, that is, to preserve milk, fruits, and vegetables, and to enhance the quality of life with the resultant beverages, cheeses, bread, pickled foods, and vinegar. The use of yeasts dates back to ancient days. The oldest fermentation know-how, the conversion of sugar to alcohol by yeasts, was used to make beer in Sumeria and Babylonia before 7000 BC. By 4000 BC, the Egyptians had discovered that carbon dioxide generated by the action of brewer's yeast could leaven bread. Ancient peoples made cheese with molds and bacteria. Wine was made in China as early as in 7000 BC [1] and in Assyria in 3500 BC. Reference to wine can be found in the Book of Genesis, where it is noted that Noah consumed a bit too much of the beverage. According to the Talmud, "a man without salt and vinegar is a lost man." The Assyrians treated chronic middle ear diseases with vinegar, and Hippocrates treated patients with it in 400 BC. According to the New Testament, vinegar was offered to Jesus on the cross. For thousands of years, moldy cheese, meat, and bread were employed in folk medicine to heal wounds. By 100 BC, ancient Rome had over 250 bakeries which were making leavened bread. As a method of preservation, milk was fermented to lactic acid to make yogurt and also converted into kefyr and kumiss using the *Kluyveromyces* species in Asia. The use of molds to saccharify rice in the Koji process dates back at least to 700 AD. By the fourteenth century AD, the distillation of alcoholic spirits from fermented

Industrial Biotechnology: Microorganisms, First Edition. Edited by Christoph Wittmann and James C. Liao.

grain, a practice thought to have originated in China or the Middle East, was common in many parts of the world. Vinegar manufacture began in Orleans, France, at the end of the fourteenth century, the surface technique being referred to as the *Orleans method*.

Antonie van Leeuwenhoek, in the Netherlands in the seventeenth century, turning his simple lens to the examination of water, decaying matter, and scrapings from his teeth, reported on the presence of tiny "animalcules," that is, moving organisms less than 1/1000th the size of a grain of sand. He was a Dutch merchant with no university training but his spare time interest was the construction of microscopes. This lack of university connection might have caused his discoveries to go unknown, had it not been for the Royal Society in England and its secretary, Henry Oldenburg, who corresponded with European science amateurs. From 1673 to 1723, Leeuwenhoek's great powers as a microscopist were communicated to the Society in a series of letters. Thus the practice of industrial biotechnology has its roots deep in antiquity.

In these early days, most scientists thought that microbes arose spontaneously from nonliving matter. What followed was an argument over spontaneous generation, aptly called the *War of the Infusions* lasting 100 years. Proponents had previously claimed that maggots were spontaneously created from decaying meat; however, this was discredited by Redi. By this time, the theory of spontaneous generation, originally postulated by Aristotle among others, was discredited with respect to higher forms of life, so the proponents concentrated their arguments on bacteria. The theory did seem to explain how a clear broth became cloudy via growth of large numbers of such "spontaneously generated microorganisms" as the broth aged. However, others believed that microorganisms only came from previously existing microbes and that their ubiquitous presence in air was the reason that they would develop in organic infusions after gaining access to these rich liquids. Three independent investigators, Charles Cagniard de la Tour of France, Theodor Schwann, and Friedrich Traugott Kützing of Germany, proposed that the products of fermentation, chiefly ethanol and carbon dioxide, were created by a microscopic form of life. This concept was bitterly opposed by the leading chemists of the period (such as Jöns Jakob Berzelius, Justus von Liebig, and Friedrich Wöhler), who believed fermentation to be strictly a chemical reaction; they maintained that the yeast in the fermentation broth was lifeless, decaying matter. Organic chemistry was flourishing at the time, and these opponents of the living microbial origin were initially quite successful in putting forth their views. Interest in the mechanisms of these fermentations resulted in later investigations by Louis Pasteur, which not only advanced microbiology as a distinct discipline, but also led to the development of vaccines and concepts of hygiene, which revolutionized the practice of medicine.

In 1850, Davaine detected rod-shaped objects in the blood of anthrax-infected sheep and was able to produce the disease in healthy sheep by inoculation of such blood. In the next 25 years, Pasteur of France and John Tyndall of Britain demolished the concept of spontaneous generation and proved that existing microbial life came from preexisting life. In the 1850s, Pasteur detected two

distinct forms of amyl alcohol, that is, D and L, able to polarize light in different directions (opticals isomers or enantiomers) but he was not able to separate the two. He found that only one of the two optical isomers (e.g., for tartaric acid) were produced by living microbes carrying out fermentation. Pasteur concluded in 1857 that fermentation was a living process of yeast. In 1861, he proved the presence of microbes in air and discredited the theory of spontaneous generation of microbes. It was at this point that microbiology was born, but it took almost two decades, until 1876, to disprove the chemical hypothesis of Berzelius, Liebig, and Wöhler, that is, that fermentation was the result of contact with decaying matter.

In 1876, the great German microbiologist, Robert Koch, proved that bacteria from anthrax infections were capable of causing the disease. His contributions involving the growth of microbes in pure culture led to the decline of the pleomorphism theory, that is, that one form of bacteria developed into another. It was mainly the work of Koch that led to the acceptance of the idea that specific diseases were caused by specific organisms, each of which had a specific form and function. In 1884, his students, Gaffky and Loeffler, were able to confirm the etiologic role of infectious bacteria in the cases of typhoid fever and diphtheria and, in 1894, Alexandre Yersin, Louis Pasteur's student, for bubonic plague. Yersin also confirmed the presence of the disease organism in the animal vector, rats.

The distillers of Lille in France called upon Pasteur to find out why the contents of their fermentation vats were turning sour. He noted through his microscope that the fermentation broth contained not only yeast cells but also bacteria that could produce lactic acid. He was able to prevent such souring by a mild heat treatment, which later became known as *pasteurization*. One of his greatest contributions was to establish that each type of fermentation was mediated by a specific microorganism. Furthermore, in a study undertaken to determine why French beer was inferior to German beer, he demonstrated the existence of strictly anaerobic life, that is, life in the absence of air. Interest in the mechanisms of these fermentations resulted in the later investigations by Pasteur, which not only advanced microbiology as a distinct discipline, but also led to the development of vaccines and concepts of hygiene, which revolutionized the practice of medicine. With the establishment of the germ theory of disease by Pasteur and Koch, the latter half of the nineteenth century was characterized by the fight against disease and the attention of microbiologists was directed to the medical and sanitation aspects of microbiology. Owing to the work of Pasteur and Koch, it became evident that the body's own defenses played a great part in fighting pathogenic microbes. It was found that when a bacterium invaded the body of a human or an animal, proteins (i.e., antibodies) were formed in the bloodstream. These could specifically neutralize the invading parasite. The science of immunology was thus founded. By injecting either dead forms or attenuated forms of the disease-producing bacterium, Pasteur could render the individual immune to the disease. The production of these vaccines occupied much of the early research in microbiology.

The application of antiseptics materialized t the time of the contributions made by Pasteur. It had been shown in 1846 by Semmelweis that chlorine could control infection, and in 1865, Joseph Lister showed that the same could be done with carbolic acid. Later, Paul Ehrlich used synthetic dyes and established the concept of the "magic bullet." Toward the end of the nineteenth century, Ehrlich began testing many synthetic compounds. He achieved success in 1909, curing relapsing fever, syphilis, and trypanosomiasis with an arsenical product called *Salvarsan* or *Compound 606*, because it was his 606th attempt to produce an arsenical compound which killed the syphilis bacterium *in vivo* without harming the host. This was the first chemotherapeutic drug ever discovered and he coined the term *chemotherapy*. This use of drugs selectively toxic to the parasite but not damaging to the host opened an entirely new field for the curing of human diseases. In 1927, this work was continued by Gerhard Domagk in Germany along with his collaborators Mietzsch and Klarer. They were working at the I.G. Farbenindustrie which was the result of a 1924 merger between Bayer and BASF. Their work resulted in the development of the red-colored molecule Prontosil rubrum. This compound was active in mice against streptococci but strangely was not active *in vitro*. Then in 1935, Trefouel and co-workers in France discovered that the red dye was broken down in the animal to the colorless and inhibitory sulfanilamide. This discovery of the first "pro-drug" also established the important concept that chemicals could kill or inhibit bacteria without toxicity to humans. Although the Nazi government refused to permit Domagk to accept the Nobel Prize in 1939, he later accepted it in 1947. Other synthetic chemotherapeutic drugs gained wide use over the years, including isonicotinic acid hydrazide and *para*-aminosalicylic acid, both for tuberculosis.

For thousands of years, moldy cheese, meat, and bread had been employed in folk medicine to heal wounds. In the 1870s, Tyndall, Pasteur, and William Roberts, a British physician, directly observed the antagonistic effects of one microorganism on another. Pasteur, with his characteristic foresight, suggested that the phenomenon might have some therapeutic potential. During the ensuing 50 years, various microbial preparations were tried as medicines, but they were either too toxic or inactive in live animals. This led to the momentous moment in microbiological history, when, in 1927, Alexander Fleming discovered penicillin (see Section 1.3).

In 1877, Moritz Traube proposed that (i) proteinlike materials catalyzed fermentation and other chemical reactions and (ii) they were not destroyed by such activities. This was the beginning of the concept of what we call enzymology today. He also proposed that fermentation was carried out via multistage reactions in which the transfer of oxygen occurred from one part of a sugar molecule to another, finally forming some oxidized compound such as carbon dioxide and a reduced compound such as alcohol. The field of biochemistry became established in 1897 when Eduard Buchner found that cell-free yeast extracts, lacking whole cells, could convert sucrose into ethanol. Thus, the views of Pasteur were modified and it became understood that fermentation could also be carried out in the absence of living cells.

During World War I, the need for glycerol, used to manufacture ammunition, resulted in the application of yeast to convert sugars into glycerol. This development led to an exhaustive study after the war of the mechanisms involved in these reactions and those converting sugars to ethanol by Neuberg. This was followed by the studies of the Dutch in Delft dealing with oxidation/reduction reactions and the kinetics of enzyme-catalyzed reactions.

Also during World War I, Chaim Weizmann of the United Kingdom applied the butyric acid bacteria, used for centuries for the retting of flax and hemp, for production of acetone and butanol. His use of *Clostridium* during World War I to produce acetone and butanol was the first nonfood fermentation developed for large-scale production; with it came the problems of viral and microbial contamination that had to be solved. Although use of this fermentation faded because it could not compete with chemical means for solvent production, it did provide a base of experience for the development of large-scale cultivation of fungi for production of citric acid. Soon after World War I, an aerobic process was devised in which *Aspergillus niger* was used (see Section 1.2). Not too many years later, the discoveries of penicillin and streptomycin and their commercial development heralded the start of the antibiotic era (see Section 1.3).

1.2
Primary Metabolites and Enzymes

1.2.1
Birth, Rise, and Decline of the Term "Biotechnology" in the Period 1900–1940

The word "biotechnology" was coined around 1919 by the Hungarian agricultural engineer Karoly Ereky, who used the term in the title of his book "Biotechnologie der Fleish-, Fett-, und Milcherzeugung im Landwirtschaflichen Grossbetriebe" (Biotechnology of meat, fat, and milk production in large-scale agricultural industries). Ereky, who later became Hungarian Minister of Food, had established a large intensive pig-rearing farm and processing plant close to Budapest, Hungary, where pigs (called *biotechnological working machines*) converted agro- and waste streams into meat, fat, and leather. In the previous decades, chemistry had merged with technology and had resulted in a novel fast-growing industry, the chemical industry. Erecky envisaged agriculture and biology combined with engineering to lead to a new industrial revolution. His vision, soon to be applied to microorganisms, rather than macroorganisms, became popular among agrobiologists, chemists, and engineers. On the basis of his perception, by fermenting cheap and abundant agricultural produce and waste, both the farmers and the chemical industry became beneficiaries. This vision led scientists and engineers to produce a range of "bio"-chemicals (solvents, alcohols, organic acids, and enzymes) using starch/sugar-fermenting microbes. Although inspirational to many scientists and engineers, his new term *biotechnology* was hardly used at all and was almost

forgotten until 1975–1980, whereas the then existing terms such as industrial fermentation and industrial microbiology remained widely used till the late 1980s [2].

1.2.2 Influential Scholars Boosting Industrial Fermentation from 1900 to 1940

In the late nineteenth century, several renowned scientists believed that the emerging industrial application of microbiology would form a new type of industry, differing from the then rapidly growing (petro)chemical industry. This idea was, at least in Europe, based on the huge importance and value of the German beer industry at the turn of the nineteenth century; it was second only to machinery building and surpassed metallurgy and coal mining. Indeed, on the basis of Pasteur's theories and practical findings in France, combined with those of Koch and Cohn in Germany, Lister in the United Kingdom, and Emil Christian Hansen in Denmark, brewing had evolved from an art into a controlled and well-understood malting, mashing, and yeast fermentation process. Also at that time, yeast culture collections were established in Prague, Delft, Berlin, and fermentation and brewing research institutes were founded (Pasteur Institute, Paris; Carlsberg Institute, Copenhagen; Institut fur Gärungsgewerbe (Institute for Fermentation Industries), Berlin). They soon gained impact and fame and still continue to function today, although under other names. In 1898, an English translation appeared of Franz Lafar's famous two-volume handbook in German, "Technical Mycology: The Utilization of Micro-organisms in the Arts and Manufactures." Lafar, the first director of the Vienna Technical Institute, became famous for his improvements of alcohol fermentation and distillery practice.

World War I brought on innovative fermentation applications. In the United Kingdom, Chaim Weizmann, who was trained at the Institut Pasteur in Paris, worked at Manchester University closely together with a brewing equipment manufacturer, R. Seligman, who had introduced the plate heat exchanger. In 1915, Weizmann developed a suitable method to ferment potato starch and grain with anaerobic bacteria to produce the chemical, acetone, on a large scale. Acetone was essential for the manufacture of much-needed ammunition for the British Army. In Germany, in 1915, W. Connstein and K. Lüdecke developed fermentation processes for glycerol, lactic acid, and yeast for animal feed under the pressure of World War I. In the 1920s, Ereky's "Biotechnologie" vision was soon applied to microorganisms (rather than to pigs) by the German microbiologist Paul Lindner, a pupil of Koch, based at the Inst. Gärungsgewerbe, Berlin. This trend was followed up especially in Czechoslovakia, The Netherlands, the United Kingdom, and the United States.

In the 1930s, at the Charles University in Prague, Prof. Konrad Bernhauer became a fervent promoter of the fermentation-based chemistry. His classic textbook of 1936, "Gärungschemisches Praktikum" (Practical Chemistry of Fermentation), condensed the knowledge of fermentation in Europe and the United States. After World War II, he became an important mentor of German scholars at the Inst. Gärungsgewerbe. Ereky and Bernhauer can be considered

the prewar advocates of industrial uses of (micro)biology; however, their Nazi links caused their names to become forgotten in later years.

In the Netherlands, at the Technical University Delft, the fermentation metabolism group of J.A. Kluyver became influential in the 1920s and 1930s in providing basic insights of microbial growth, metabolism, and production potential. In 1921, Kluyver became the Chair of General and Applied Microbiology, upon Martinus Beijerinck's retirement. In 1924, he investigated the production of sorbose by *Acetobacter suboxydans* and collaborated with the "Nederlandse Gist-en Spiritus Fabriek" (The Dutch Yeast and Alcohol Manufacturing Company). Over the next few years, he described chemical transformations performed with microbes in a scientific way, including oxidations, fermentations, and incomplete oxidations. By 1926, he had published his famous paper, "Unity and Diversity in the Metabolism of Microorganisms," and explained the term *facultative anaerobes*. He also made a valuable and industrially relevant contribution by developing the technique of submerged culture of molds, later to be widely used in the fermentation industry. His PhD student, C.B. Van Niel discovered and developed the aroma compound of butter, that is, diacetyl, important as a bioflavor for the growing margarine industry. Van Niel left Delft toward the end of 1928 to accept an offer to become Professor at Stanford University's Hopkins Marine Station, in Pacific Grove, CA, USA. His research on photosynthetic bacteria revolutionized the concept of the biochemistry of photosynthesis.

In the United States, especially at the US Department of Agriculture (USDA), employing about 600 chemists in 1915, research boosted the industrial use of biology, especially due to the dislocation of chemical supplies as a result of World War I. The dairy chemist, James Currie, worked on the production of citric acid with the mold *A. niger*. He persuaded, in 1923, the then small, New York-based Chas. Pfizer and Co. to support him. He developed the surface fermentation process in shallow trays to convert sugar into citric acid, which until then, had to be extracted from lemons and other citrus fruits. In 1929, Pfizer switched to submerged fermentation based on the research of Bernhauer. Also in the 1930s, well before World War II, the Research Director of the Dow Chemical company, William J. Hale, promoted heavily the use of chemicals, including ethanol (called agricrude-alcohol), made from cheap farm produce. This principle was named "Chemurgy" by him. He advocated the creation of "Agricenters" for processing of farm products into industrial end products and of raw materials for other process industries. In the meantime, USDA researchers in Washington, DC, developed microbial processes for the production of other organic acids from sugar and starch (i.e., lactic acid, gluconic acid, and others) for use in the food and other industries. However, their laboratory was abandoned to make space for construction of the Pentagon. Four new regional laboratories were set up, including the Northern Regional Research Laboratories (NRRL) in Peoria, Illinois during 1939–1940. It was there that on July 14, 1941, Florey and Heatley arrived from Oxford University with Fleming's penicillin fungus *Penicillium notatum* in their coat pockets! (see Section 1.3).

These developments in biotechnology, during the 1900–1930s, occurred along with those in petrochemical engineering to form a novel and separate field of

science and technology. The German chemical companies (e.g., Bayer, BASF, and Hoechst) and several oil companies (BP, Shell, and Standard Oil) were set up to become established firms, although they were not as important then as they became later. In retrospect, the oil crises in 1973 and 1979 forced the chemical and oil industries again to reorientate and this also boosted renewed interest in Ereky's "Biotechnologie" and Hale's "Chemurgy" concepts. The term *industrial biotechnology* surfaced again in the 1980s.

1.2.3 Milestone Achievements in Industrial Fermentation Technology

1.2.3.1 The Acetone–Butanol–Ethanol (ABE) Fermentation Process

By the start of the twentieth century, shortages of natural rubber activated interest in alternative feedstocks and in chemical routes to produce synthetic rubber. This attracted the attention of the young chemist, Chaim Weizmann, who was assistant to Prof. W.H. Perkin at Manchester University, UK. The chemical company Strange and Graham Ltd (London) had also shown interest in a process to prepare butadiene or isoprene, building blocks of rubber by oxidation of *n*-butanol or isoamylalcohol, both obtainable by fermentation of sugars. They recruited Perkin and Weizmann to work on this project. This joint research project had to be refocused during World War I, because Britain's need for acetone as a solvent for the manufacturing of smokeless explosive cordite became critical. Butanol and acetone had already been reported as fermentation products by Pasteur in 1861 and F. Schardinger in 1905. Weizmann was able to select a superior strain of *Clostridium acetobutylicum*, which produced commercially interesting levels of acetone, butanol, and ethanol (ABE) using cereals as feedstock, and he filed a patent in 1915. Owing to the German blockade, Britain soon experienced a shortage of grain and decided to move the solvent production plants to Canada and India. Also in the United States, the US Air Service and the British War Mission purchased the Commercial and Majestic Whiskey distilleries in Terre Haute, Indiana, and modified them for acetone production using the Weizmann process. The Commercial Solvents Corp. of New York managed the new company. Between May and November 1918, 400 000 l of acetone were produced with 800 000 l of *n*-butanol as a coproduct. These surpluses of *n*-butanol became valuable during the prohibition era in the United States (1920–1933), as it could replace amylacetate in lacquers (for automobiles). Butanol also found use in solvents, plasticizers, paints, and resins. From the 1930s onward, the butanol fermentation process was largely superceded by its petrochemical production route. However, today, it is again gaining commercial interest. Weizmann's research work on acetone and its essential role in the British war period was recognized by the Cabinet Minister of Armament David Lloyd George, who later became Prime Minister. Weizmann, always a fervent proponent of a homeland for the Jewish people, was to later (1948) become the first President of the State

of Israel. After 1940, continued interest in solvent fermentations [3, 4] led to the further development of the butanol fermentation [5].

1.2.3.2

A Novel Vitamin C Fermentation Process

Early observations on microbial oxidations of sugar alcohols (polyols) culminated in a novel process for vitamin C (ascorbic acid) in the 1930s. In 1867, Pasteur had observed that certain bacteria, which he called *Mycoderma aceti*, oxidized the alcohol in wine into acetic acid (to make vinegar). In 1886, in the United Kingdom, Adrian J. Brown used *Bacterium aceti* (now *Acetobacter aceti* subsp. *xylinum*) to oxidize mannitol to fructose, *n*-propanol to propionic acid, and ethyleneglycol to glycolic acid. In 1898, G. Bertrand reported on the microbial oxidation of other polyols to ketones, for example, sorbitol to sorbose, using Brown's strain; this also laid the basis for the Bertrand – Hudson rule. Revisiting the work of Bertrand in the early 1930s, Tadeus Reichstein from the Chemistry Department of ETH in Zurich, Switzerland, successfully devised a microbial approach for oxidizing D-sorbitol to L-sorbose, an important intermediate in the chemical synthesis of vitamin C. This bioconversion step worked so efficiently that the company F. Hoffmann-La Roche AG in Basel decided to produce vitamin C via this chemoenzymatic route, rather than extracting it from fruits. They used *A. suboxydans* cultures to convert 20% solutions of D-sorbitol into L-sorbose with yields of up to 97%.

Today, various combinations of chemical and microbiological approaches are still used to meet the high demand for vitamin C as a nutriceutical and an antioxidant [6–8]. Other vitamin processes important today include riboflavin [9–12] and vitamin B12 [13, 14].

1.2.3.3

The Lactic Acid Fermentation Process

In 1857, Pasteur described what he called a *lactic yeast*, responsible for the formation of lactic acid, when advising a distillery experiencing difficulties in the fermentation of sugar beet juice to ethanol. Lactic acid remained a specialty product until 1883, when the young MIT-educated chemist Charles E. Avery built the first lactic acid fermentation plant, the Avery Lactate Company, in Littleton, near Boston, Massachusetts; the fermentation substrate used was hydrolyzed corn starch. Avery's aim was to replace cream of tartar (potassium bitartrate) used as an acidulant in the bakery sector. This project was initially successful but after a fire ruined the plant in 1911, several other US companies replaced Avery's company in this effort. Competition also came from lactic acid producers in Germany, for example, Boehringer Co., Knab and Lindenhayn and E. Merck. They switched later to whey, molasses, or sugar as substrate [15]. During World War I, lactic acid production increased considerably in Germany, to meet the military requirements to replace glycerol Also in the United Kingdom and France, new facilities for the lactic acid fermentation were built.

In the United States, the group of L.A. Rogers at the USDA's Bureau of Dairy Industry introduced the use of pure cultures in the American dairy industry

and devised a continuous fermentation process for lactic acid based on whey. In 1936, based on the work of the Rogers group, large-scale operation was realized by Sheffield By-Products Company, Norwich, NY, using wooden fermentation vats to grow *Lactobacillus bulgaricus* and stainless steel equipment to counteract the corrosive properties of lactic acid. Also, American Maize Products Co., DuPont, and Clinton Corn Syrup Refining Co. started to produce lactic acid from glucose-rich starch hydrolysates and corn steep liquor in the late 1930s. Then, in Europe, several new lactic acid plants became operational, for example, Byk-Guldenwerke and C. H. Boehringer Sohn (Germany), Société Normande de Produits Chimiques (France), Bowmans Chemicals (UK), Schiedamsche Melkzuur Fabriek (SMF, later named CCA) (The Netherlands), and Kemisk Vaerk Koge (KVK) (Denmark). Other companies were established in Italy, Hungary, Czechoslovakia, Poland, Romania, and Russia. Also, five companies in Japan (Takeda Chemical Industries, Tanabe, Sankyo, Dai Ichi Seiyaku, Dai Nippon Industries) producing lactic acid were operational in 1937. Synthetic lactic acid, manufactured first by Musashino Chem. Lab. Ltd, Tokyo, then came on the market. Petrochemical technology was also increasingly applied to produce ethanol, acetone, and butanol. However, new applications were developed for lactic acid and its derivatives in the food industry, and the medical, health and technical sectors. This led to a revival of the lactic acid fermentation process [16, 17] with several new companies that are still active today.

1.2.3.4 Fermentative Production of Glycerol

In Germany, during World War I, factories focused on manufacturing glycerol by fermentation, equally needed for their weapons and explosives industry. This was based on the "Protol" process, developed in 1915 by W. Connstein and K. Ludecke. They had found that addition of sodium bisulfite to a yeast ethanol fermentation process using beet sugar diverted it into a process yielding glycerol. In 1919, Carl Neuberg and co-worker J. Hirsh revealed the mechanism of bisulfite action, that is, it prevented alcohol formation and rerouted toward glycerol overproduction. For some time, 24 factories in Germany produced 12 000 tons of glycerol per year for use in the production of explosives. Today, glycerol is available in large quantities as a valuable side product of biodiesel production and of fat hydrolysis.

1.2.3.5 L-(−)-Ephedrine by Fermentation

In 1921, Neuberg and Hirsch discovered that yeasts could condense added benzaldehyde with pyruvate-derived acetaldehyde to form the chiral product, L-(+) phenylacetylcarbinol (also named Neuberg's ketone). This alpha-hydroxyketone (acyloin) can easily be chemically converted into L-(−)-ephedrine, an important bronchodilator still made by a bioprocess.

1.2.3.6

Steroid Transformations

An old paper by Lintner and von Liebig in 1911 on the reduction of furfural to furfurol by yeast attracted the interest in 1937 of Mamoli and Vercellone, former students of C. Neuberg. It inspired them to use yeast to reduce 4-androstenedione to testosterone. This was the first example of a successful microbial steroid transformation, to be followed by many more in the early 1950s.

1.2.3.7

The Citric Acid Fermentation Process

Citric acid was commercially produced from the 1820s until about 1919 from Italian lemons; then, microbial citric acid took over. In 1893, Carl Wehmer, while at the Technical College in Hannover, Germany, became interested in mycology and studied fungal metabolic acids, oxalic acid, and citric acid. Two fungal species of a genus which he called *Citromyces (=Penicillium)*, were able to produce considerable levels of citric acid when grown on the surface of 10% sugar solutions. He recognized the importance of his findings and applied for patents in 1894. However, owing to technical and sterility issues, his project never got beyond the pilot scale. Several years later, in 1917, James N. Currie, a dairy scientist at the USDA, who was aware of Wehmer's papers, examined several other molds. He discovered *A. niger* to be a good producer of citric acid when cultured in media with low pH, high sugar levels, and mineral salts. He informed Chas. Pfizer & Co. Inc., then a major producer of lemon-derived citric acid, of his findings and requested commercial interest toward his microbial process. He was then hired by the company and asked to develop his findings into a commercial process. This formed the basis of the first citric acid plant in the United States, in 1923. The Pfizer plant dominated the citric acid market for many years to come and they also built an overseas plant in the United Kingdom in 1936. Similar industrial processes had started in Belgium (in 1919) and led to companies such as the companies Citrique Belge S.A., John and E. Sturge in the United Kingdom, Montanindustrie J. D. Starck A.G. in Czechoslovakia, and Boehringer and J.A. Benckiser in Germany; they all used the "surface culture" process. Details are not well documented owing to the restriction of information by the manufacturers. The fungal mycelium was grown as a surface mat on liquid beet molasses medium in a large number of shallow trays, stacked in a large room kept under semiaseptic conditions; spores were blown in with a sterile air stream. This process was profitable for many years and was only challenged in the 1940s by the development of submerged fermentation processes for citric acid. Today, the latter is still the main process for fermentative production of citric acid [18, 19] as well as other organic acids [20, 21], including acetic acid [22–25].

1.2.3.8

Gluconic Acid Process

Formation of gluconic acid was first observed by Boutroux in 1880 using the bacterium *M. aceti* b (*A. aceti*). In 1922, Molliard described formation of gluconic

acid by the mold *A. niger*, along with citric acid and oxalic acid. A few years later (1924), Bernhauer found an *A. niger* strain that almost exclusively formed gluconic acid, when grown as thin mats on glucose solutions at low temperature. Over the coming decades, this fermentation process was intensively studied and optimized by researchers at the USDA and in Japan, where surface as well as submerged fungal fermentations under increased air pressure and at high glucose levels (up to 35%) were developed. Today, such processes are used for large-scale gluconic acid production. 2-Keto-gluconic acid is also produced by fermentation [26].

1.2.3.9

Other Important Fermentation Processes and Products

Other bacterial-based fermentations such as 2,3-butanediol, acetoin, dihydroxyacetone, keto-gluconate, propionic acid, vinegar, and "old traditional" fermented foods (e.g., cheese, yoghurt, pickles, and sauerkraut) were studied during this period both to gain more basic microbiological and biochemical understanding, as well as to develop large-scale controlled fermentations. During the period 1900–1930, important "traditional" yeast-based fermentations, such as the production of food, baker's and feed yeasts, beer brewing, beverage, industrial, and fuel alcohol, were further optimized and reached high volumes worldwide. From the 1930s onward, industrial and fuel alcohols were increasingly produced by chemical synthesis from petroleum feedstock. In the United States in 1936, about 84% of ethanol was still produced by fermentation of different agro-derived substrates (molasses, grain, sulfite liquor, etc.), while only 16% was made from ethyl sulfate via chemical synthesis; in 1946, the figures changed to 64% versus 36%, respectively. This chemical synthesis trend continued for a while to overtake the use of fermentation. However, owing to high petroleum prices and environmental concerns, the tide turned and industrial and fuel ethanol also began to be made microbiologically [27–31]. Other fermentations based on fungal strains, including itaconic acid, kojic acid, fumaric acid, and gallic acid, have been studied by several research groups in Japan, the United States, and in Europe. Their industrial production became very important after World War II. Also of great significance were the fermentations developed in the late 1900s for amino acids [32] especially those for L-glutamic acid [33–37] and L-lysine, as well as those devised for 5′ nucleotides such as guanylic (GMP) and inosinic (IMP) acids [38–40]. Fermentative production of polymers such as dextran, xanthan [41], polyhydroxy butyrate [42, 43], and polylactic acid (PLA) [44] also became important.

1.2.3.10

Applied Biocatalysis and Industrial Enzymes

Although several practical developments in the field of biocatalysis date from the first half of the nineteenth century (e.g., use of diastase extracted from malted barley in the brewing industry) and Emil Christian Hansen's enzyme preparation, rennet, for cheese making (1874), scientific background on enzymes

only emerged later in the nineteenth century. This was based on the findings of Emil Fisher starting in 1894 on enzyme specificity and its "lock and key" action and on the 1897 work of Eduard and Hans Buchner on the pure chemical nature of the alcohol fermentation in the absence of living yeast cells. The soluble agent in yeast press juice was called *zymase*. The work eliminated the "*vis vitalis*" (vital force) paradigm altogether. A further key step toward the "chemical paradigm" was the work of J.B. Sumner in 1926 on the crystallization of jack bean (*Canavalia ensiformis*) urease and on the protein nature of enzymes. In the 1930s, several more enzymes were isolated, purified, and crystallized from plants, animal organs, as well as from yeasts, molds, and bacteria. Technical developments on enzymes started at the onset of the twentieth century with the founding of the Rohm and Haas Company in 1907 in Germany, and the description of several practical enzymatic reactions with crude amylase, lipase, protease, trypsin, pepsin, invertase, and others. The kinetic studies by Michaelis and Menten in 1913 were also very important toward the understanding of the physicochemical nature of enzyme action. The Japanese scientist Jokichi Takamine, working in the United States (Peoria, Illinois), was the first to patent a microbial enzyme product (1894). This "Takamine" process involved extraction with aqueous ethanol of extracellular amylases (named "Taka-diastase") from *Aspergillus oryzae*, growing on bran (similar to the ancient Japanese koji process). Early in the twentieth century, plant lipases were produced by mechanical disruption of ricinus seeds and used to produce fatty acids from oils and fats. It was also found that this reaction is reversible and the enzymatic synthesis of fat from glycerol and fatty acids was described as early as in 1911. Proteolytic enzymes were successfully used in 1911 in the United States for the chillproofing of beer. Wheat diastase was found to interact beneficially with dough making and the addition of malt extract became a common practice in bread baking. Production of pectinases started in Europe in the 1930s for use in the fruit juice sector. For leather manufacturing, early tanners kept the animal skins in a warm suspension of dog and bird dung, not knowing that this unpleasant bating practice was based on the action of enzymes (pepsin, trypsin, lipase, etc.) present in animal dung. Once this mechanism was revealed in 1898, a bacterial bate was developed from *Bacillus erodiens* cultures and commercialized as a bacterial culture ("Erodin") adsorbed on wood meal. In 1907, pancreatic extract was introduced as a bating agent by O. Rohm, who founded his own company in Stuttgart, Germany. With the trade name "Oropon," his product became very successful and he moved production to larger facilities in Darmstadt. Here, a growing market, searching for a new and pleasant technical product, was an important factor in his success. It also led to the increasing knowledge on the principles of enzymatic action. Further development of large-scale submerged fermentation processes for enzymes has led to increased industrial production and applications of enzymes. This happened in the late 1950s, with the emergence of detergent enzymes and use of glucoamylase to produce glucose from starch.

1.3 The Antibiotic Era

1.3.1 Penicillin

The very first recorded observation on microbial "antibiosis" dates back to 1877 [45]. Pasteur and Joubert described slower growth of *Clostridium* sp., in the presence of other bacteria. In 1893, Bartolomeo Gosio, an Italian physician, discovered a compound in the culture filtrate of *Penicillium brevicompactum*, which, in pure crystallized form, inhibited growth of *Bacillus anthracis*; it was later rediscovered and named mycophenolic acid [46]. Although it was never used as an antibiotic, owing to its toxicity, a derivative found use as a new immunosuppressant. In the early 1920s, André Gratia, a microbiologist at the University of Liége, Belgium, studied the lysis of bacteria by products derived from other microorganisms. He was one of the first phage researchers after F. d'Herelle, belonging to the period before the viral nature of bacteriophages became clear. In 1925, Gratia described the bacteriolytic effect of certain fungi, including a *Penicillium* strain that exerted this action on anthrax-causing bacteria. Owing to an illness, Gratia did not further pursue this research topic [47].

The accidental discovery of penicillin by Alexander Fleming in 1929 in England began the golden era of antibiotics. He noted that some of his plates containing *Staphylococcus aureus* were contaminated with the mold, *P. notatum*. Strangely, he observed that none of the bacterial colonies could grow in the vicinity of the mold and concluded that the mold was producing some inhibitory agent. He also noted that filtrates of the mold lyzed the staphylococci and were nontoxic in animals. Because of his earlier discovery and studies on lysozyme, he recognized this as an important phenomenon to pursue. He coined the name penicillin for the antibacterial substance in the mold culture broth, and published his findings in 1929. Since the activity was very unstable and Fleming could get no encouragement from his fellow scientists concerning the usefulness of such material, the project was abandoned by Fleming.

Although Fleming's discovery led to penicillin, the first successful chemotherapeutic agent produced by a microbe, thus initiating the golden age of the wonder drugs, the road to development of penicillin as a successful drug was not an easy one. Attempts to isolate penicillin were made in the 1930s by a number of British chemists, but the instability of the substance frustrated their efforts. For a decade, penicillin remained a laboratory curiosity. With the advent of World War II, and the deaths of many British soldiers on the battlefield from bacterial infections after being wounded, a study of penicillin began in 1939 at the Sir William Dunn School of Pathology of the University of Oxford by Howard W. Florey, Ernst B. Chain, Norman G. Heatley, Edward Abraham, and their colleagues. Chain had searched through the papers of Fleming and others on lysozyme, penicillin, and metabolites with inhibitory effects on other organisms.

In the years 1939–1941, under eminent war pressure, Fleming's *P. notatum* strain was grown as a surface culture in large flasks, fermentation media were optimized, penicillin activity measurements were improved, solvent extraction of penicillin from the broth was optimized, toxicity tests in animals (mice) were conducted, as well as protection tests by penicillin of mice infected with lethal doses of virulent *Streptococcus* sp. This amazing effort led to the successful preparation of a stable form of penicillin and the demonstration of its remarkable antibacterial activity and lack of toxicity in mice. Subsequent clinical trials on humans were very successful and time had come for commercial production. Production of penicillin by the strain of *P. notatum* in use was so slow, however, that it took over a year to accumulate enough material for a clinical test on humans; large-scale production became essential. Since British pharmaceutical companies did not show interest, the Oxford University administration was forced to contact the Rockefeller Foundation in New York that had funded their research. Florey and Heatley were sent to New York in 1941 and met with R.G. Harrison, Chair of the National Research Council, who advised them to contact the USDA. There they met Percy A. Wells, acting Chief of the Bureau of Agriculture and Industrial Chemistry. He was in charge of the four regional research laboratories and, fortunately was a fermentation specialist. He sent Florey and Heatley to the USDA's NRRLs in Peoria, Illinois, resulting in a historical outcome.

Florey and Heatley convinced the NRRL and several American pharmaceutical companies (including Merck, Squibb, and Pfizer), to develop the production of penicillin. Heatley remained for a period at the NRRL to work with Moyer and Coghill and then worked for a while at Merck. Thus began a momentous cooperative effort among university and industrial laboratories in the United States and academic institutions in England, which lasted throughout the war.

Although Fleming's original strain produced only traces of penicillin, "brute force" genetic manipulation made tremendous strides in production ability and led to a whole new technology known as *strain improvement*. These early basic genetic studies concentrated heavily on the production of mutants and the study of their properties. The ease with which "permanent" characteristics of microorganisms could be changed by mutation and the simplicity of the mutation technique had tremendous appeal to microbiologists [48]. A cooperative "strain-selection" program was established between workers at the USDA in Peoria, the Carnegie Institution at Cold Spring Harbor in New York, Stanford University and the University of Wisconsin. Strain selection began with *Penicillium chrysogenum* NRRL-1951, the well-known isolate from a moldy cantaloupe obtained in a Peoria market. This strain was capable of producing 60 $\mu g\,ml^{-1}$. Cultivation of spontaneous sector mutants and single-spore isolations led to higher-producing cultures from NRRL 1951. One of these, NRRL 1951-1325, produced 150 $\mu g\,ml^{-1}$. It was next subjected to X-ray treatment by Demerec of the Carnegie Institution, and the mutant X-1612 was obtained, yieldeding 300 $\mu g\,ml^{-1}$. Workers at the University of Wisconsin obtained ultraviolet-induced mutants of Demerec's strain. One of these, Q-176, which produced 550 $\mu g\,ml^{-1}$, became the ancestor of all of the strains subsequently used in industry. The

"Wisconsin family" of superior strains became well known all over the world, some producing over 1800 $\mu g\,ml^{-1}$. The penicillin improvement effort was the start of a long engagement between genetics and industrial microbiology, which ultimately demonstrated that mutation was the major factor involved in the 100- to 1000-fold increases obtained in production of microbial metabolites.

Thousands of lives were saved, on and off the battlefield. This discovery and development of the β-lactam antibiotics was among the most powerful and successful achievements of modern science and technology. Since Fleming's accidental discovery of the penicillin-producing mold, years of steady progress followed, and today, the β-lactam group of compounds is the most successful example of natural product application and chemotherapy.

The discovery and development of the sulfa drugs and penicillin have been reviewed by Dixon [49] and Bentley [46].

Penicillin had been originally produced in surface culture, but titers were very low. Submerged culture soon became the method of choice. By the use of strain improvement and medium modifications, such as the use of corn steep liquor as additive, the yield of penicillin was increased by 100-fold in just a few years. Much of the understanding of the physiology of *P. chrysogenum*, in relation to penicillin production, was achieved by Professor Marvin Johnson and his students at the University of Wisconsin. Further clinical successes were demonstrated in both England and the United States, and finally in 1943, penicillin was used to treat those wounded in battle.

By the 1950s, it was realized the *P. chrysogenum* could use additional acyl compounds as side-chain precursors (other than phenylacetic acid for penicillin G) and produce new penicillins, and one of these, penicillin V (phenoxymethylpenicillin), achieved commercial success. Its commercial application resulted from its stability toward acid which permitted oral administration, an advantage it held over the accepted article of commerce, penicillin G (benzylpenicillin). Penicillin G and penicillin V became the main penicillins of commerce. In commercial production, the usual medium had been a complex one, composed of glucose, corn steep liquor, side-chain precursor (phenylacetic acid for penicillin G or phenoxyacetic acid for penicillin V), and mineral salts. The earliest recognition that glucose had a negative effect on penicillin biosynthesis was made by Johnson and his students. They found glucose to be excellent for growth but poor for penicillin formation, while lactose showed the opposite pattern. They devised a medium containing both sugars in which growth occurred at the expense of glucose and when it was exhausted, the mass of cells began to produce the antibiotic at the expense of lactose. Unlike glucose, lactose was utilized slowly and did not exert carbon catabolite repression on the process. Then, Davey and Johnson found that intermittent or continuous feeding of the less expensive glucose could replace batch feeding of lactose. This represented the birth of the method of fed-batch fermentation which is commonplace in the fermentation industry today.

The biosynthesis of penicillin from its precursors, L-cysteine, L-valine, and phenylacetic acid, was actively studied during the 1950s, 1960s, and 1970s. Of great interest was the relationship between L-lysine and penicillin formation. In

1947, David Bonner made the observation that 25% of the lysine auxotrophs that he had made from *P. chrysogenum* failed to make penicillin and he predicted that (i) there was some relationship between the antibiotic and the amino acid and (ii) there was a common precursor of the two compounds. That he was absolutely correct was established later when Demain and co-workers found that lysine was a potent inhibitor of penicillin biosynthesis [50, 51]. That the inhibition could be reversed by L-α-aminoadipic acid led to the postulations that (i) L-α-aminoadipic acid was involved in penicillin biosynthesis although it did not end up in the final penicillin molecule, (ii) penicillin was derived from α-ketoglutarate and acetyl-CoA via the fungal lysine biosynthetic pathway, and (iii) lysine inhibition of penicillin biosynthesis was due to feedback inhibition by lysine of its own biosynthetic pathway, thus limiting L-α-aminoadipic acid formation. Independently, Arnstein and colleagues detected the tripeptide α-aminoadipyl-cysteinylvaline (ACV) as an intracellular compound in *P. chrysogenum*. Results in several laboratories established L-α-aminoadipic acid as an important precursor of all penicillins. Soon, α-(L-aminoadipyl)-L-cysteinyl-D-valine (LLD-ACV) was established as the crucial intermediate of penicillin biosynthesis. The reaction, sensitive to feedback inhibition by lysine, was later proven to be the initial step of lysine biosynthesis in fungi, that is, the homocitrate synthase reaction.

The future of penicillins became doubtful in the 1950s as resistant strains of *S. aureus* emerged in hospital populations. The staphylococcal population was building up resistance to penicillin via selection of penicillinase-producing strains, and new drugs were clearly needed to combat these resistant forms. Also, the penicillins described up to this point were solvent-soluble, exhibiting a high degree of activity against gram-positive organisms but were much less active against gram-negatives. Fortunately, two developments occurred which led to a rebirth of interest in the penicillins and related antibiotics. One was the 1959 discovery in Japan by Koichi Kato of the accumulation of the "penicillin nucleus" in *P. chrysogenum* broths to which no side-chain precursor had been added. In England, Batchelor and co-workers isolated the compound (6-aminopenicillanic acid, 6-APA) which was the "penicillin nucleus" discovered by Kato. 6-APA was later used to make "semisynthetic" (i.e., chemical modification of a natural antibiotic) penicillins with the beneficial properties of resistance to penicillinase and to acid, plus broad-spectrum antibacterial activity.

The second development was the discovery of a completely different type of penicillin, a hydrophilic type, showing equivalent activity against both classes of microorganisms (i.e., penicillin N). This compound was discovered independently by two groups of workers. Brotzu in 1948 published his work in an unknown Sardinian journal on the isolation of an antibiotic-producing culture of the fungus *Cephalosporium acremonium* (later reclassified as *Acremonium chrysogenum*) from sewage. Unable to purify the antibiotic material, he sent the culture to Florey at Oxford where some 10 years earlier, *P. notatum* had been at a similar stage in history. While the British were studying the components of this antibiotic complex, workers at the Michigan Department of Health announced that a species of *Tilachlidium* produced a new antibiotic which they called *synnematin*.

After the culture was reclassified as *Cephalosporium salmosynnematum*, it was shown that synnematin was a complex of two components, A and B. Almost nothing was ever published on the A component, and we have no knowledge today about its structural relationship to synnematin B. While the above work was going on in Michigan, the British announced that Brotzu's strain produced two antibiotics, that is, "cephalosporin P," active only against gram-positive organisms and "cephalosporin N," which was active against both gram-positive and gram-negative bacteria. "Cephalosporin P" was found to be of steroidal nature and not a β-lactam at all. "Cephalosporin N," on the other hand, was found by Abraham and co-workers to be a true penicillin possessing an α-D-aminoadipyl side chain, and to be identical to synnematin B. It was renamed "penicillin N." In comparison to penicillin G, it was only 1% as active against gram-positive forms, but had equal to or somewhat greater activity against gram-negative bacteria. The hydrophilic nature of penicillin N and its roughly equivalent activity against gram-positive and gram-negative bacteria were due to the carboxyl group in the side-chain.

The tremendous success attained in the battle against disease with penicillin G not only led to the Nobel Prize being awarded to Fleming, Florey, and Chain, but to a new field of antibiotic research, and a new antibiotic industry. Penicillin opened the way for the development of many other antibiotics, and it still remains the most active and one of the least toxic of these compounds. Today, more than 100 antibiotics are used to combat infections to humans, animals, and plants.

1.3.2 The Cephalosporins

A key breakthrough was the finding in Edward Abraham's laboratory at Oxford mentioned above, that is, that a second antibiotic was produced by Brotzu's strain of *A. chrysogenum*. After his earlier contributions as part of the Florey penicillin team, Abraham had established an independent laboratory at Oxford. Abraham and Newton found the new compound to be related to penicillin N in that it consisted of a β-lactam ring attached to a side chain which was identical to that of penicillin N, that is, D-α-aminoadipic acid. It differed, however, from the penicillins in containing a six-membered dihydrothiazine ring in place of the five-membered thiazolidine ring of the penicillins. It was called *cephalosporin C*. Thus, the era of the cephalosporins was launched.

The nucleus of cephalosporin C was named 7-aminocephalosporanic acid (7-ACA). Cephalosporin C strongly absorbed ultraviolet light, was stable to acid and to penicillin β-lactamase, was non-toxic and had *in vivo* activity in mice. Its mode of action was the same as that of the penicillins; that is, inhibition of bacterial cell wall formation. Although neither penicillin N nor cephalosporin C was ever commercialized, they led to important knowledge on the biosynthesis of these compounds and the development of many powerful semi-synthetic cephalosporins of great use in medicine.

The stability of cephalosporin C to penicillin β-lactamase was a very attractive property. The main disadvantage of the molecule was its weak activity; it had only 0.1% of the activity of penicillin G against sensitive staphylococci, although its activity against gram-negative bacteria equaled that of penicillin G. However, by chemical removal of its D-α-aminoadipic acid side chain and replacement with phenylacetic acid, a penicillinase-resistant semi-synthetic compound was obtained which was 100 times more active than cephalosporin C. Many other new cephalosporins with wide antibacterial spectra were developed in the ensuing years, for example, cephalothin, cephaloridine, and cephaloglycin, making the semi-synthetic cephalosporins the most important group of antibiotics at that time. The stability of the cephalosporins to penicillins was evidently a function of its dihydrothiazine ring because (i) the D-α-aminoadipic acid side chain did not render penicillin N immune to attack and (ii) removal of the acetoxy group from cephalosporin C did not decrease its stability to penicillinase. Cephalosporin C competitively inhibited the action of penicillinase from *Bacillus cereus* on penicillin G. Although it did not have a similar effect on the *S. aureus* enzyme, certain of its derivatives did. Another major advantage was that cephalosporins could be given to some patients who were allergic to penicillins.

From a biosynthetic sense, the relationship of penicillin N and cephalosporin C was of great interest. An important development that led to rapid progress in this area was the subcellular work done by Abraham and his Oxford colleagues in the 1970s [52]. They used protoplast lysates of *A. chrysogenum* to convert labeled valine into a penicillin. This led to the discovery by Kohsaka and Demain at MIT in 1976 [53] of the ring expansion reaction, catalyzed by the "expandase" (deacetoxycephalosporin C synthase; DAOCS) enzyme. For many years, it had been thought that penicillin N and cephalosporin C were products of different biosynthetic branches in *A. chrysogenum*. However, the discovery of the ring expansion enzyme showed that cephalosporins were produced from a penicillin. This was confirmed by Yoshida *et al.* [54] who showed that the ring expansion enzyme converted penicillin N to deacetoxycephalosporin C. Extracts of mutants which fermentatively produced penicillin N but not cephalosporins failed to carry out this reaction, whereas early blocked mutants (negative for production of both penicillin N and cephalosporins) did expand the ring.

From the 1970s through the 1980s, knowledge accumulated concerning fermentation and biosynthesis of cephalosporin C. The following were of major importance: the (i) stimulation by DL-methionine via a regulatory mechanism unrelated to its ability to contribute the sulfur atom to the antibiotic, (ii) use of acetate as precursor to the acetoxy group, (iii) L-cysteine and L-valine as precursors of the nucleus, and (iv) L-α-aminoadipic acid as precursor of the D-α-aminoadipyl side chain of cephalosporin C. An important step forward was provided by Banko and co-workers [55] when they demonstrated that the cell-free activity of the enzyme from *A. chrysogenum* formed the important tripeptide precursor of all penicillins and cephalosporins, that is, LLD-ACV. This enzyme, ACV synthetase, was proven to be a single multifunctional enzyme acting on L-α-aminoadipic acid, L-cysteine,

and L-valine to produce LLD-ACV. Also of importance was the isolation by Hollander and co-workers [56] of pure isopenicillin N synthetase ("cyclase") which converted the LLD-ACV to isopenicillin N.

The long-held notion that β-lactams were produced only by fungi was shattered by a report from Merck & Co. that a streptomycete produced penicillin N. This provocative finding was presented at the 1962 American Society for Microbiology Annual Meeting and was only published as an abstract. Although much doubt was cast on this report, two reports from Eli Lilly & Co. [57] and Merck & Co. [58] published some 9–10 years later reported that various species of *Streptomyces* and *Nocardia* produced cephalosporins modified at C7 (=cephamycins) and/or at the side chain attached to C3. The discovery of cephamycin C led to much research on and development of prokaryotic cephalosporins as the presence of the methoxy group on the β-lactam ring made the molecule more active against gram-negative and anaerobic pathogens and more resistant to gram-negative β-lactamases. For the first time in the history of the β-lactams, molecules were available which showed a high degree of stability toward these troublesome enzymes. As with fungal cephalosporin C, cephamycin C was never used clinically but was employed for semisynthesis of many medically useful compounds. A more potent semisynthetic cephamycin, cefoxitin, was rapidly commercialized by Merck, to be followed later by cefmetazole, temocillin, cefotetan, and other semisynthetic cepahlosporins.

In the 1970s and the 1980s, the pathways to the penicillins and the cephalosporins including cephamycin C were worked out especially after cell-free systems became available [59]. Late in the 1970s came reports on the production of β-lactam antibiotics which were neither penicillins nor cephalosporins. The most important was clavulanic acid from streptomycetes, which possessed only weak antibiotic activity but was an excellent inhibitor of β-lactamase [60]. It became a blockbuster compound by being coformulated with broad-spectrum semisynthetic penicillins which are susceptible to β-lactamase, for example, with amoxycillin, the combination is known as *Augmentin*.

Another important development in the history of the β-lactam antibiotics was the discovery of the carbapenems at Merck. The first, called *thienamycin*, was discovered by Kahan and his group with a screening protocol based on inhibition of peptidoglycan synthesis [61]. The antibiotic was produced by *Streptomyces cattleya*, which also made cephamycin C. Carbapenems resembled the penicillins in having a β-lactam ring fused to a five-membered ring. They differed in that the five-membered ring was unsaturated and contained a carbon atom instead of the sulfur. Sulfur was, however, present in another location in all the carbapenems produced by streptomycetes. A large number of carbapenems were reported, but thienamycin was the most important. Indeed, it was the most potent, most broad-spectrum, and nontoxic natural antibacterial agent ever found. It inhibited cell wall synthesis, as did the penicillins and cephalosporins, and was relatively resistant to microbial β-lactamases.

The development of commercial process technology for the penicillin nucleus (6-APA) and the cephalosporin nucleus (7-ACA) opened the way for chemical

acylation with various side chains yielding many improved semisynthetic penicillins and cephalosporins with broader antibacterial activity and improved pharmacokinetic properties. Broad-spectrum penicillins and cephalosporins became the best-selling antibacterial agents in the pharmaceutical arena.

1.3.3 The Waksman Era

The advent of penicillin, which signaled the beginning of the antibiotic era, was closely followed in the 1940s by the discoveries made by Selman A. Waksman, a soil microbiologist at Rutgers University. He and his students, especially H. Boyd Woodruff, Albert Schatz, and Hubert Lechevalier, succeeded in discovering many new antibiotics from the filamentous bacteria, the actinomycetes, such as actinomycin, neomycin, and the best-known of these new "wonder drugs," streptomycin. These discoveries on the antibiotic-producing abilities of the actinomycetes occurred long before the developments, described above, of β-lactams produced by these filamentous bacteria. Waksman and Woodruff published in 1940 on the discovery of the actinomycins, which were chromo-oligopeptides. One such compound, actinomycin D, has been used for years to combat the Wilms tumor in children and became a very important tool in the development of molecular biology as an inhibitor of RNA polymerase.

After its momentous discovery in 1944 by Waksman, Schatz, and Bugie as a product of *Streptomyces griseus*, streptomycin was used against tuberculosis caused by *Mycobacterium tuberculosis* and also against gram-negative bacteria; bacterial meningitis was also treated with streptomycin. Its major impact on medicine was recognized by the award of the Nobel Prize to Waksman in 1952. As the first commercially successful antibiotic produced by an actinomycete, this aminoglycoside led the way to the recognition of these organisms as the most prolific producers of antibiotics. Streptomycin also provided a valuable tool for studying cell function. After a period of time during which it was thought to act by altering permeability, its interference with protein synthesis was recognized as its primary effect. Its interaction with ribosomes provided much information on their structure and function; it not only inhibited their action but also caused misreading of the genetic code and was required for the function of ribosomes in streptomycin-dependent mutants. With Lechevalier, Waksman reported on the discovery of neomycin in 1948 and candicidin in 1953. Neomycin, an aminoglycoside produced by *Streptomyces fradiae*, served as a topical antibacterial, and the polyene candicidin, made by *S. griseus*, found use as a topical antifungal antibiotic.

Cooperation on the development of industrial processes between Rutgers University, Princeton University, Columbia University, and Merck & Co., Inc. led to the birth of the field of biochemical engineering. With royalties on streptomycin turned over to Rutgers University by the manufacturer, Merck, Waksman was able to build the world-famed Institute of Microbiology.

The discoveries of the aminoglycosides at Rutgers further advanced the antibiotic era and resulted in the later discovery of many more "wonder drugs" such as chloramphenicol in 1947, tetracyclines in 1948, macrolides such as erythromycin in 1952 [62], glycopeptides such as vancomycin in 1956, additional aminoglycosides such as gentamicin in 1963, β-lactams such as cephamycins in 1970, and carbapenems in 1979, ansamycins such as rifamycin in 1957, and polyene macrolides such as nystatin in 1950. Of 15 000 microbial secondary metabolites discovered, about 12 000 were antibiotics. Their unusual chemical structures included β-lactam rings, cyclic peptides containing "unnatural" and nonprotein amino acids, unusual sugars and nucleosides, polyenes, and large macrolide rings. Although most were useless for humans, being either too toxic or inactive in higher organisms, others were lifesavers. The antibiotics were virtually the only drugs utilized for chemotherapy against pathogenic microorganisms and were crucial in the increase in average life expectancy in the United States from 47 years in 1900 to 74 for males and 80 for women in 2000.

For some reason, the actinomycetes were amazingly prolific in the number of antibiotics they could produce. Roughly 70% of all antibiotics were obtained from these filamentous prokaryotes, and 75% of those were in turn made by a single genus, *Streptomyces*. It is quite amazing that strains of *Streptomyces hygroscopicus* produce over 180 different secondary metabolites. About 10% of the antibiotics are made by unicellular bacteria and about 20% by fungi. New bioactive products from microbes were discovered at an amazing pace: 200–300 per year in the late 1970s, increasing to 500 per year by the 1990s. Accompanying the natural product antibiotics in the pharmaceutical arena were synthetic antimicrobials such as the quinolones and fluoroquinolones. Even these synthetics traced their discovery back to a natural product, that is, quinine. The first quinolone, nalidixic acid, was modeled after quinine. Commercialization of antibiotics, however, slowed down in the 1980s and only three, daptomycin, caspofungin acetate, and the synthetic oxazolidinone were commercialized in the ensuing decades.

1.3.4 Mutagenesis and Strain Improvement

The experiences described above in which penicillin-producing cultures were mutagenized and screened for improved producers led the way to the extensive use of genetics to improve production capability. From the 1950s on, genetic manipulation such as mutagenesis/screening was replaced by mutagenesis/selection/screening in which various selective means were used to decrease the number of strains that had to be screened for improved production. Then, it was found that new derivatives, some better than the parent molecule, could be produced by mutants. This was first discovered by Kelner in 1949 but the more active derivatives were not isolated and identified. However, the medically useful metabolites demethyltetracycline and doxorubicin were discovered later by mutation of the cultures producing tetracycline and daunomycin, respectively. In 1969, University of Illinois Professors Kenneth Rinehart and David Gottlieb,

and student W.T. Shier devised and reported the technique of "mutational biosynthesis" (=mutasynthesis) [63]. In this process, a mutant blocked in secondary metabolism was fed analogs of the moiety whose biosynthesis was blocked. If successful, the mutant (called an *idiotroph*) produced a new secondary metabolite. Mutational biosynthesis was used for the discovery of many new secondary metabolites. The most well known was the commercial anthelmintic agent doramectin, production of which employed a mutant of the avermectin producer, *Streptomyces avermitilis* [64].

For strain improvement, genetic recombination was virtually ignored in industry before 1975, mainly owing to the low frequency of recombination, as low as 10^{-6}. However, use of polyethyleneglycol-mediated protoplast fusion in actinomycetes by Okanishi *et al.* [65] changed the situation markedly. Okanishi's work on protoplast formation, fusion, and regeneration accelerated the use of genetic recombination. From then on, there was a heightened interest in the application of genetic recombination to the production of important microbial products. Frequencies of recombination increased to even greater than 10^{-1} in some cases. After 1985, many strain improvement programs routinely included (i) transposition mutagenesis, (ii) targeted deletions and duplications by genetic engineering, and (iii) genetic recombination by protoplast fusion and plasmid transformation. Much was known about genetics and regulation in the actinomycetes owing to the elegant research on *Streptomyces coelicolor* by David Hopwood, Keith Chater, Mervyn Bibb, and their colleagues at the John Innes Institute in Norwich, England (which has been called the Temple of Streptomyces Genetics) [66]. Their favorite organism made at least five secondary metabolites (an A-factor-like molecule, and the antibiotics, actinorhodin, undecylprodigiosin, methylenomycin A, and "calcium-dependent antibiotic" or CDA).

These efforts and others by geneticists in academia and industry throughout the world in the 1970s and 1980s revealed that the genes encoding most antibiotic biosynthetic pathways were clustered into operons, thus facilitating transfer of entire pathways from one organism to another. Such clusters were found to also include regulatory and resistance genes. In 1985, "combinatorial biosynthesis" was born. An international effort from the United Kingdom, Japan, and the United States resulted in the cloning of a pathway from one streptomycete producing the isochromanequinone antibiotic actinorhodin into strains producing granaticin, dihydrogranaticin, and mederomycin (which are also isochromanequinones). This resulted in the discovery of two new hybrid antibiotics, mederrhodin A and dihydrogranatirhodin [67]. Combinational biosynthesis became a widely used technique used for discovery of new hybrid drugs [68] by recombinant DNA (rDNA) technology. New antibiotics were discovered [69–71] including those created by changing the order of the genes of an individual pathway in its native host.

Progress in strain development has recently involved extensive use of new genetic techniques such as (i) metabolic engineering, accomplishing quantification and control of metabolic fluxes, and including inverse metabolic engineering and transcript expression analyses such as association analysis and massive parallel signature sequencing; (ii) directed evolution; (iii) molecular breeding

including DNA shuffling and whole genome shuffling; and (iv) combinatorial biosynthesis. These efforts facilitated not only the isolation of improved strains but also the elucidation and identification of new genetic targets to be used in product discovery.

1.3.5 Semisynthetic Antibiotics to Combat Resistant Microbes

Although there was a feeling expressed by many in the late 1970s that the era of product discovery for bacterial disease was ending, the battle against resistant microbes continued and featured some surprising developments. These included the following: (i) semisynthetic variations of old antibiotics (ketolides, clarithromycin, azithromycin, glycylglycines); (ii) older underutilized antibiotics (teicoplanin); (iii) new derivatives of previously undeveloped narrow-spectrum antibiotics (streptogrammins); and (iv) a few newly developed antimicrobial agents (caspofungin, daptomycin, and the synthetic epothilones).

Of great success was the development of semisynthetic erythromycins. These included clarithromycin, roxithromycin, azithromycin, and the ketolide, telithromycin. While the first two showed improved acid stability and bioavailability over erythromycin A, they showed no improvement against resistant strains. On the other hand, azithromycin and telithromycin acted against macrolide-resistant bacteria. All the above semisynthetic erythromycins were effective agents for upper respiratory tract infections and could be administered parentally or orally. Telithromycin was bacteriostatic, active orally, and of great importance for community-acquired respiratory infections. Of particular interest was its low abilities to select for resistance mutations and to induce cross-resistance. It also did not induce Macrolide-Lincosamide-Streptogramin B-resistance (MLSB), a problem with other macrolides. For more than 35 years, the glycopeptides vancomycin and teicoplanin were virtually the only natural antibiotics active against multidrug-resistant gram-positive bacteria. Their use was cut down by the increase in multidrug resistance. To the rescue came a number of new semisynthetic antibiotics, one called Synercid. Synercid was composed of a synergistic (by 100-fold) pair of narrow spectrum streptogramins, that is, quinupristin and dalfopristin, both being semisynthesized from natural compounds made by a single strain of *Streptomyces pristinaespiralis*. The pair was constituted by a (Group A) polyunsaturated macrolactone containing an unusual oxazole ring and a dienylamide fragment and a (Group B) cyclic hexadepsipeptide possessing a 3-hydroxypicolinoyl exocyclic fragment. Although the natural streptogramins were poorly water soluble and could not be used intravenously, the Synercid components were both water soluble. They inhibited protein synthesis and were active against vancomycin-resistant *Enterococcus faecium* (VREF), methicillin-resistant *S. aureus* (MRSA), glycopeptide-resistant *S. aureus*, and β-lactam-resistant *Streptococcus pneumoniae*. Synergistic action of the streptogramins was due to the fact that the B component blocked binding of aminoacyl-tRNA complexes to the ribosome, while the A component inhibited peptide bond formation and

distorted the ribosome, promoting the binding of the B component. Synercid was approved by FDA in 1999. A semisynthetic tetracycline, that is, a glycylcycline, was successfully developed for use against tetracycline-resistant bacteria. The 9-*t*-butylglycylamido derivative of minocycline called *tigecycline*, is active against resistant gram-positive, gram-negative, and anaerobic bacteria possessing the ribosomal protection resistance mechanism or the active efflux mechanism.

In addition to antibiotics, many other extremely important secondary metabolites were discovered or developed over the years, such as hypocholesterolemic agents including compactin, mevinolin, pravastatin; enzyme inhibitors, such as lipstatin, clavulanic acid, polyoxins, and siderophores; immunosuppressants, for example, sirolimus (a rapamycin), cyclosporine A, and tacrolimus (FK-506); antitumor agents, such as taxol and ergot alkaloids; bioherbicides, including biolaphos and phosphinothricin; plant growth stimulants, such as the gibberellins; antiparasitic agents, including monensin, polyethers, and avermectins; and bioinsecticides.

1.4
The Biotechnology Era Between 1970 and 2015

1.4.1
Biotechnology in the 1960s and 1970s; Governmental and Political Initiatives

By the 1950s, large-scale production of products such as beer, cheese, citric acid, amino acids, enzymes, and pharmaceuticals such as antibiotics had become well established. Growing recognition of economic relevance followed the success of penicillin and the manufacture of other antibiotics based on applied microbiology and biochemical engineering. The most important products (commodities) were ethanol, starch products, L-glutamate, and other amino acids, citric acid, and other organic acids [72]. Enzyme application became a large, important technology, and big business, with a turnover of €2.5 billion in 2010 [73].

Starting in the 1970s and 1980s, biotechnology attracted the attention of government agencies in Germany, the United Kingdom, Japan, the United States, and other countries as a field with innovative potential, capable of contributing to economic growth. A first enthusiastic report by the German chemical technology organization DECHEMA was produced in 1974 for the German Ministry for Education and Science (Bundesministerium für Bildung und Wissenschaft, BMBW). It was the first systematic approach for biotechnology research funding, emphasizing classical biotechnology and aiming to develop a unified research and development strategy. The German government wanted to accelerate biotechnology research and development to identify and encourage innovations in industry [2, 74–76]. This study has been an intriguing example of interaction between politicians, industrialists, and scientists and was termed a *corporatist approach* by Jasanoff [77]. Subsequent studies on biotechnology in other European countries were initiated, for example, in the United Kingdom, Japan, and France [2, 74].

Most topics, however, were classical or conventional. The focus of interest was on the identification of new products and new or improved products, processes, and services. rDNA methods were not mentioned, as they were still in early development, and there was no example that this was a truly enabling technology at the time of the study (1972–1974). Interdisciplinary communication between microbiologists, chemists, and engineers was difficult owing to different working procedures, specialized scientific and technical languages, and different approaches to this new field. No integration of methods and theoretical approaches took place or was developed seriously. The major industries in Germany, as well as in the rest of Europe, notably the chemical, pharmaceutical, and food industries, were conservative in their attitude toward the new field of biotechnology. University education essentially remained confined within traditional disciplines, faculties, and curricula. Exceptions to this were the specialized research units at a few UK, European, and American universities that offered special courses (e.g., University College London and the Technical University of Berlin). The first biotechnology journal of high reputation was established in 1958 by Elmer Gaden, "Journal of Microbiological and Biochemical Engineering." It later became "Biotechnology and Bioengineering." A few other journals were launched in the 1950s and 1960s (such as "Applied Microbiology," renamed "Environmental and Applied Microbiology," and "Applied Microbiology and Biotechnology").

In contrast to Europe, in the United States, during the 1970s and subsequently in the 1980s, a more profound change took place including innovation based on recombinant technologies. This is evidenced notably by the emergence of biotechnology companies during the 1970s and 1980s and by an OTA (Office of Technology Assessment, USA) study [78]. The emphasis in the OTA study was on genetic engineering and rDNA technology resulting in commercial opportunities and support for the rapid commercial exploitation of scientific results; it was closely associated with the business world: "This report focuses on the industrial use of rDNA, cell fusion, and novel bioprocessing techniques. In the past ten years, dramatic new developments in the ability to select and manipulate genetic material have sparked unprecedented interest in the industrial uses of living organisms." Emphasis was placed on the pharmaceutical sector, which had been most active in commercializing biotechnology. Thus, the first recombinant biotechnology products, such as rDNA-produced human insulin, interferon (IFN), and monoclonal antibody (mAb) diagnostic kits, were a direct result of the basic research that led to these new technologies. The highlights of the US competitive position were a well-developed life-science base, the availability of financing for high-risk ventures, and an entrepreneurial spirit that had led the United States to the forefront of biotechnology commercialization. The political framework made it possible for industrialists and scientists to rapidly capitalize on the results of basic research. The transfer of science to the economic sphere resulted in the continued founding of many new biotechnology companies; and the development and marketing of a considerable range of new important products, many blockbusters, mostly pharmaceuticals which were previously not accessible. It was a type of gold rush to support the "New Biotechnology," as recombinant technology was

known in the United States. The story of insulin (see the following) is an example of both a great breakthrough and the delayed recognition of the potential of genetics and recombinant technologies by established pharmaceutical companies. As an anecdotal example of the reticence prevailing in European Industry at that time, John Collins, in 1975 and later, Herbert Boyer together with Robert Swanson (founders of Genentech) presented their positive viewpoint of the impact of rDNA-genetics on the synthesis of human insulin at the Novo Company in Denmark, which supplied 30% of the insulin world market. Their views were taken seriously by young scientists but, initially, not by management as they believed that there was no possible solution to the problems involved in the industrial development of recombinant bacteria.

1.4.2 The Development of New Biotechnology Companies

The turning point in genetics ensued from the establishment of a model for the molecular structure of DNA by James Watson and Francis Crick, based on the crystallography data of Rosalind Franklin. But the "*DNA Revolution*" as Hotchkiss termed it, progressed or penetrated slowly into technology, initially having little effect on traditional processes and products [79]. Significant scientific breakthrough events and technological progress provided a new basis for BT (see Section 1.5). New methods and tools that played a key role in the expansion of recombinant technologies include analysis of DNA, RNA, and proteins, and of their structure, synthesis of short DNA molecules, the identification and purification of DNA molecules that code for pharmaceutically active proteins, introduction of such DNA (also from human sources) into bacteria, and expression of the protein in bacteria.

The story begins in 1972 with the concepts for cloning – transfer of DNA from different sources into bacteria in 1972 (see Section 1.5 for details).

The spark of invention has been vividly remembered by Cohen [80] himself, Boyer, Falkow [81], and two others with the remarkable event when, during an evening in an Hawain Deli, during a conference, the concept for cloning was conceived and written on a napkin: The experiments in the laboratories of both Boyer and Cohen during the following month were successful and published in 1973 in the Proceedings of the National Academy of Sciences USA (PNAS). "Boyer's restriction enzyme (*Eco* RI) became the workhorse of the recombinant DNA revolution." It was given by Boyer to other labs working in the field ([82], p. 59). Among the difficulties was screening at the time, but Boyer's team had been able to identify the few bacterial colonies – literally one in a million – containing the toad DNA. "Herb just said he kissed every colony on the plate, until one turned into a prince" (Falkow, cited by Hall [82], p. 63). This was the basis for the first patent on recombinant technologies by Cohen and Boyer, filed in 1974, acknowledged in 1980. (For the controversial debate on patenting, see Section 1.5).

The key to the founding and the success stories of new BT companies, and the expansion of recombinant technologies, were new tools primarily based on recombinant technologies (for details, see Section 1.5).

The patent office in 1980 acknowledged that the claims by Cohen and Boyer, for the use of plasmids and restriction nucleases constituted a novel invention of recombinant-DNA technology. Stanford's decision to patent the rDNA technique and the breadths of the claims upheld by the Patent Office were widely noted, and they set patterns, both in academia and in industry [83].

In the controversy about patenting biological issues, the pro arguments were that biotechnology companies provided the most effective means of translating advances in science and technology into products for medicine and agriculture [82–84] and, of course, there were the precedents of Pasteur's patent on pure cultures of yeast and Chakrabarty's patent on multiplasmid-containing strains able to degrade xenobiotics (made by classical microbiology techniques, not rDNA) [72].

One of the questions asked in Science and Technology Studies (STS) is how basic science leads to technology and how this is converted into innovation. Biology, unlike engineering, physics, and chemistry had no tradition of basic scientists being routinely involved in innovation. Consequently, the founding of Genentech by Boyer and Swanson, the first biotechnology company based on gene cloning, in 1976, has become a paradigm for just such a transition. Swanson, a venture capitalist, and Boyer, a microbiologist and biochemist, had the joint vision to commercialize rDNA. After ups and downs, they admirably achieved their goal, an example to be followed by many others, and a trend that revolutionized modern medicine. Genentech at the outset had the immediate goal of producing pharmaceutical products and exploiting the capacity of bacteria for the production of hormones, such as human insulin and human growth hormone, which could be put to immediate clinical use.

It is to be noted that the first modern biotechnology company, Cetus, was founded in Berkeley, California in 1972, producing mainly chemicals for industry and agriculture and employing about 250 co-workers in 1980 [85]. Cetus was based on a different type of "cloning technology," namely, on the biophysics of scanning and deflecting droplets (originally the Coulter counter), so that single cell-based colonies (called clones) could be plated on growth media. This latter was combined with microscopic scanning of a large grid, a technology also used later on a microscale for high-throughput DNA sequencing, so that the automated screening of millions of single mutant clones could be carried out. Cetus after it failed owing to a delay in FDA approval of their rDNA product interleukin-2 was later incorporated into the Chiron Corporation (Berkeley, 1991). At Cetus, Kary Mullis succeeded with a highly important innovation, the polymerase chain reaction (PCR) method for amplification of DNA (even from tiny amounts). As a potent analytical tool, it had a tremendous influence on various fields of modern biotechnology, even for the identification of criminals as developed by Alec Jeffreys (University of Leicester, UK). Mullis was awarded the Nobel Prize in 1993 and remains the only employee rather than founder of

a biotechnology company to win a Nobel Prize. PCR, in turn, led directly to the large-scale development of the biochip industry by the Chiron spin-off Affymetrix (for the development of microarrays and DNA-CHIPs see [72], section 11.4.3) and now allows isolation of any gene from a known genome sequence.

Boyer and Swanson conceived Genentech, the first gene cloning company, in 1976, when they decided to commercialize rDNA. It was founded in the same year, with the immediate goal of producing pharmaceutical products, and to exploit the capacity of bacteria for the production of hormones, such as human insulin and human growth hormone, which could be put to immediate clinical use. The story will be presented in some detail, as it was this success that finally convinced industry that the new biotechnology was really an approach to be taken seriously, as, in particular, the final product was approved for clinical use in a surprisingly short time after the gene had been cloned and expressed (4 years project development time).

Boyer at Genentech and Arthur Riggs and Keiichi Itakura from the Beckman Research Institute, were the first to successfully express a human gene (for somatostatin, a hypothalamus hormone used for several hormonal disorders, previously made by synthesis) in bacteria in 1977 [72]. Somatostatin's small size made it particularly appealing to Riggs and Itakura to synthesize the gene for it. It could be interesting in itself (although necessary for a relatively rare disorder, gigantism). For Swanson, validating the development model was necessary, that is, providing the gene for a protein (e.g., a hormone) and cloning and expressing it in bacteria. The immediate plan was to go for insulin, in order to enter a big market [82]. In addition, cooperation agreements were made with the University of California, San Francisco (UCSF) and the City of Hope National Medical Center, California. The company strategy comprised to make as many patents as possible, to produce scientific publications in high-ranking journals, and to present results at international meetings, a strategy that should, and did, attract many brilliant young scientists.

Involved in the race for insulin, and the "DNA gold rush," were some of the most outstanding scientists in molecular biology, namely, Walter Gilbert of Harvard, Howard Goodman and William Rutter of UCSF, Boyer with David Goeddel, Axel Ullrich, and Peter Seeburg of Genentech [82, 83]. Goeddel, described as a "kamikaze scientist" by colleagues became the first full-time Genentech employee, followed, with some delay, owing to their commitment to their academic careers as postdoctoral fellows at UCSF, by Axel Ullrich, Peter Seeburg (both from Germany) and John Shine (from Australia). They became a formidable, highly qualified, and engaged staff trained in rDNA cloning. Ullrich was the first to clone rat insulin, while Seeburg succeeded in cloning the rat growth hormone at UCSF, but they had been excluded as inventors on patents. In early 1980, the Genentech staff already amounted to 110 employees, 80 of whom were scientists. The years, 1978–1982, were golden ones in the Genentech laboratories, during which bright minded spirits, working furiously, "won many of the cloning races" [83]. The "mad race" won by Goeddel and Ullrich at Genentech over Walter Gilbert at Harvard, for cloning the gene for human insulin, has been recounted engagingly by Hall [82].

Genentech helped to develop 7 of the 13 types of biotechnology-based biopharmaceuticals on the market in 1993; its scientists produced 250 papers annually and 1200 granted patents by 1993. The first important contract was with AB Kabi (Sweden) for the production of somatostatin [85].

The first product commercially manufactured using gene technology was human insulin, which was approved for clinical use in 1982. A connection between Genentech and the Eli Lilly and Company was crucial. It provided Genentech financial assistance, improved the primitive Genentech expression systems to achieve commercially viable levels, provided the manufacturing facilities for a qualitative drug, and contributed the regulatory expertise, which, in 1982, brought the first rDNA drug to the market [83]. For this development, it was important that the Supreme Court of the United States allowed the patenting of new recombinant bacteria [85].

This marked the beginning of the new rDNA era of biotechnology. It now became possible to produce human proteins, hormones, IFNs, interleukins (ILs), antibodies, and so on, for medical treatment of many diseases, a tremendous advance which had not been possible earlier. It created a new "big business," with the introduction of many blockbuster pharmaceuticals. The demonstration that approval and worldwide marketing of insulin could be achieved in only 4 years after the insulin gene had been cloned belied the considerable skepticism that prevailed within the pharmaceutical industry. In particular, the demonstration that a small bioreactor of some 10 l was sufficient to provide enough material for the entire world market of human growth hormone was a dramatic step and went a long way in stimulating enthusiasm for investment in the new area of biotechnology [72] (see Table 1.3; Section 1.4.4.2).

The following are among the impressive list of innovative recombinant medicinals produced by Genentech (later, Genentech/Roche): human growth hormone (1985), IFN α-2a (1980) (in cooperation with Hoffmann-La Roche, Switzerland), Protropin (1985), Activase (tPA) (touted to be the blockbuster drug of the decade) (1987), Rituxan (1997), and Herceptin (1998); and, in addition, Thymosin α_1, proinsulin, and leucocyte and fibroplast IFNs (U. Behrendt, 2009, personal communication) [85, 86]. Scaling up was important, when, in 1979, Genentech approached the rDNA Advisory Committee for permission to use 750-l-reactors, which was granted.

The founding of Genentech had been achieved with $100 000, which was extended by $950 000 in 1978. In the initial 1980 public stock offering, Genentech raised $35 million. For tPA, Genentech had mobilized a huge force to make and market the drug during the early 1980s; however, it was no longer the house of science. Genentech was taken over by Hoffman La-Roche AG, (CH), in 1990, with a majority share of 60% for $2.1 billion, with full control and integration into the company in 2009 ([83], pp. 200–202).

Other selected biotechnological companies are listed below in chronological order.

Genex Corp. Rockville, Maryland, USA was founded 1977, with a main interest in chemicals for industry and IFN [85].

Biogen S.A., was founded in 1978 by a group of leading molecular biologists in Geneva, Switzerland, among them Heinz Schaller and Walter Fiers. It was first run by Dan Adams, and from 1982 by Walter Gilbert. The company first succeeded in the production of an IFN (Avonex) based on the research of Charles Weismann, Zurich. It generated revenues via licenses to other pharmaceutical companies, for example, Schering Plough (Intron A), and via a number of hepatitis B-based vaccines sold by Smith-Kline Beecham and by Merck [82, 83, 85, 87].

Amgen (named for Applied Molecular Genetics), another success story, was started in 1980 in Thousand Oaks, California on the initiative of a venture capitalist, William K. Bowes, who convinced Winston Salser of the Moleculat Biology Institute at UCLA to participate. Salser immediately set out to assemble a distinguished scientific advisory board, including Eugene Goldwasser of the University of Chicago, creator of valuable small amounts of erythropoietin (EPO). The diverse interests of the scientific advisors were reflected in the many directions Amgen took in the early years, apparently confused and without adequate resources, in contrast to Genentech. The leadership desperately needed finally came when George Rathman accepted the presidency. He started with private financing of $19 million, by far the largest initial equity financing in biotechnology history. He then succeeded in early 1983 to raise $43 million via public offering. An early investment by Schering Plough (which thus entered the biotechnology field) was also important. Amgen succeeded in the development of the first two blockbuster drugs for patients with blood disorders or patients undergoing chemotherapy: (i) Epogen (EPO, erythropoetin) for which the market grew to unforeseen dimensions, with excellent clinical responses in 1987 and (ii) Neupogen (G-CSF, granylocyte colony-stimulating factor), entering clinical trials, also in 1987. Amgen also developed Neulasta, as well as an interleukin-1 (IL-1) receptor antagonist. Two more rounds of financing raised another $150 million. In 2002, Amgen acquired Immunex Corporation, another biopharmaceutical company. By mid-2006, seven of its recombinant products had been approved [83, 87].

Chiron Corp., Emeryville, California, founded in 1978 by William J. Rutter, pursued, unlike other start-ups, a wide menu of pharmaceuticals, including vaccines, diagnostics, and therapeutics. Hepatitis Delta (HDV) was cloned and characterized in Michael Houghton's laboratory in 1986 and in 1989 also Hepatitis C virus (HCV) [72, 83]. This latter along with PCR-based diagnostics allowed much-improved (1000-fold increased sensitivity; also for HIV) safety for products still derived from blood serum, for example, for serum albumen used to stabilize other rDNA products used intravenously.

Large pharmaceutical companies entered the field of recombinant drugs later than the new biotechnology companies. They either acquired new biotechnology companies or established in-house know-how, or combined both strategies. Conservative as they had been, many neglected the potential of recombinant technologies, and lost their original strength in several cases, even undergoing a significant decline ([72], section 17.6.3). Remarkable changes in the structure

of the pharmaceutical industry occurred during the last century, with the trend toward "big pharma," mainly in the United States and Britain, by mergers and acquisitions (globalization) in the late 1990s and subsequent years and by the emergence of the new biotechnology companies. Data from Ernst & Young indicate that over the past few years, biotechnology firms have become the primary source of approved new medical entities (NMEs). In 2005, major pharmaceutical companies garnered only 11 NME approvals, whereas biotechnology firms had 18, with research budgets of only about one-quarter the size of that of the pharmaceutical industry [88a,b]. In 2007, the world market for pharmaceuticals was over $600 billion. A list of the largest biotechnology companies and their sales is shown in Table 1.1, with sales.

A most important market is diagnostics, with sales of about $35 billion, dominated by Roche with $9.2 billion in 2008, followed by Abbot, Johnson & Johnson, and Bayer [90, 91]. In one important segment, notably diagnostics, enzymes and mABs products are developed by rDNA tools, pioneered from the early 1980s by Boehringer Mannheim (Germany) (taken over by Roche in 1997). The large-scale production of the first recombinant enzyme was achieved in 1981, as well as an enzyme immunoassay for the identification of potato viruses, and, in 1982, test kits for human diagnosis and food analysis with recombinant enzymes.

The strong performance of the core biotechnology firms, that is, small- or medium-sized enterprises, in drug discovery has been of major importance. A boost in the founding of new companies followed during the 1980s and 1990s,

Table 1.1 (a) The largest pharmaceutical companies, with sales in 2014 and (b) the largest biotechnology companies with sales listed [72, 89a,b].

(a) Major pharmaceutical companies (selling biopharmaceuticals)	Sales ($ billion) 2014	(b) Major biotechnology companies	Sales ($ billion) 2014 (except other data given)
Johnson & Johnson (USA)	74.3	Gilead Sciences Inc. (USA)	24.9
Pfizer (USA)	49.6	Amgen Inc. (USA)	20.0
Roche (CH)	48.0	Novo (DK)	12.9
Sanofi-Aventis (F)	41.0	Genentech (USA)[a)]	10.5 (2008)
Merck & Co. (USA)	42.2	Biogen Idec Inc. (USA)	9.7
AbbVie (USA)[b)]	20.0	Celgene Corp. (USA)	7.7
Eli Lilly (USA)	19.6	Genzyme Corp. [c)] (USA)	4.0 (2010)
Total Big Pharma	430.5	Serono [d)] (CH)	2.8 (2006)
		Chiron (USA) [e)]	1.9 (2006)
		Total sales of large biotech companies	62

New data for 2014 from C&EN 2015, March 9, pp. 25, 26.

a) Taken over by Hoffman La-Roche AG, CH, in 1990, with a majority share of 60%, with full control and integrated into the company in 2009.

b) AbbVie formed by split of former Abbot Laboratories in 2013.

c) Taken over by Sanofi in 2011.

d) Taken over by Merck KGA, Germany, in 2006.

e) Taken over by Novartis, CH, in 2005/2006.

first in the United States, and later in Europe. Start-up companies were founded often on a key technology platform. In 2000, there were around 1273 new biotechnology companies with 162 000 employees and $21 billion turnover in United States. About the same time, there were 1570 new biotechnology companies with 61 000 employees and $7.7 billion turnover in Europe (data from and analysis of companies with a maximum of 500 employees in 2000) [92]. Economic data for the biotechnology firms (not including major pharmaceutical companies) have been assembled by Lähteenmäki and Lawrence [88a,b]. Biotechnology companies, whose shares are traded publicly on stock exchanges, that is, 309 firms, generated almost $47 billion in revenue in 2004; 404 firms generated more than $63 billion in 2005 in the United States.

Recently, a remarkable series of new companies, four within 4 years, have been founded in order to exploit a new tool, a technology to explore genomics and genome editing tools of remarkable precision; two founded by the inventors Jennifer Doudna and Emmanuelle Charpentier. These companies created an investment of $345 million within that short time. However, in the area of personalized medicine, some have ethical concerns with respect to germline alterations (as opposed to somatic tissue alterations; see also the end of Section 1.5) [93–95].

1.4.3 New Bioengineering Tools

The key to the founding and the success stories of new biotechnology companies, and the expansion of their technologies, were new tools, primarily based on recombinant technologies. These are discussed in detail in Section 1.5.

Classical biochemical engineering had the following aims [96, 97]: (i) the quantitative investigation of biotransformations and modeling of the processes; (ii) the development of bioreactors and downstream operations; and (iii) the transformation of laboratory results and theoretical approaches (models) into technical dimensions (scale-up). Sustainability (sustainable development) became another relevant aspect, which has been defined as development that meets the needs of the present without compromising the ability of future generations to meet their own needs. Others define it as the optimal growth path that maintains economic development while protecting the environment, optimizing social conditions, and relying on limited, exhaustible natural resources.

Classical fermentative production of products such as ethanol, amino acids, and antibiotics made significant technical progress, including feeding, mixing systems, and process analytics from the 1950s until the 1970s. The essential technologies for production are fermentation, cell culture technology, media design and preparation, downstream processing, filtration (membrane filtration, virus filtration), membrane chromatography, containment technologies, including single-use processes (single-use bags), fermenters, and downstream technologies [72].

The clinical and commercial success that recombinant proteins have had since the mid-1990s, clearly stimulated development of mammalian cell culture technology (for details, see [72], chapter 15 and [98]). The first recombinant therapeutic protein (tissue plasminogen activator, tPA) made in cultivated mammalian cells obtained market approval in 1986 (see Table 1.3), and made Chinese hamster ovary (CHO) cells the dominant host system for manufacturing more than 60% of all new target proteins in the clinical pipeline. Rapid and often dramatic yield improvements were achieved from a titer range of $5-50\,mg\,l^{-1}$ in the early 1980s to $0.5-5\,g\,l^{-1}$ in 2002–2004, an increase by two orders of magnitude. This made it possible to produce complex recombinant proteins for clinical application in kilogram quantities, or even up to a ton per year. The products exhibit all the necessary secondary (post-translational) modifications that only a higher eukaryote can execute, and that are often essential for biological activity and pharmacokinetics: proper folding, disulfide bridge formation, oligomerization, proteolytic processing, phosphorylation, and glycosylation, among others.

Reactor technology remains diversified, with reactor types ranging from roller bottles to stacked plates, hollow fibers, stirred-tank reactors, or disposables such as single-use bioreactors (SUBs), notably for low volume and speciality applications such as the production of viral vaccines [99]. Both CHO and NSO (mouse myeloma derived) cells grow well in suspension with high cell densities. Modern production systems use stirred-tank systems (U. Behrendt, 2009, personal communication) [86]. SUBs are used in the 15–125 l range for seeding larger stainless steel reactors, particularly for mammalian cell culture.

Another important topic is downstream processing. The production of biopharmaceuticals requires the highest standards of sterility and purity, including GMP [72]. Downstream requirements specific for rDNA products include, for example, the renaturation of "inclusion bodies" – proteins aggregated in bacterial cytoplasm.

Bioengineering developed toward biosystems engineering. With the progress of molecular biology, new tools were developed, often referred to as the *omics*, that is, genomics, transcriptomics, proteomics, metabolomics, fluxomics, and more [72]. Bioinformatics has been developed in order to handle, process, and correlate the huge amount of data generated by ever-faster analytical procedures. In order to combine them, a "holistic" approach known as *biosystems engineering* was created. Its aim comprises the quantitative description and improvement of established or developing novel production processes. When searching for a new product or process, it has been proposed that the ideal strategy to follow is a rational procedure of sequencing genomes of microbial species (bacteria, yeasts, or molds) of particular interest for use in metabolic engineering [100]. The next step is annotation and reconstruction of the metabolism. Transcriptome, proteome, and metabolome investigations are essential steps as the basis for rational optimization of a microorganism. Regulatory control of genes is an essential aspect to be considered. Biosystems engineering must take into account the environment of the cell in the bioreactor, namely, concentrations

(and their gradients) of substrates, notably oxygen, products, temperature, and pH, that is, the integration of physiology and fluid dynamics in bioprocesses. A further step should address the integration of all steps, from genetic engineering to downstream operations right from the beginning of process development [101]. The integration of all data using bioinformatic tools should enable the construction of networks and models using the genome and further experimental data obtained, both for production processes and target identification for drug design [100]. The analytical methods and tools used include DNA-arrays in transcriptomics (with up to 15 000 oligonucleotides); 2D-gel electrophoresis for qualitative analysis, ELISA for obtaining quantitative data in proteomics; MFA (metabolic flux analysis via mass balances and including isotope materials), FBA (flux balance analysis via modeling) in fluxomics; GC-MS in metabolomics, and classical analysis of substrates, products, pH, pO_2, CO_2, temperature, time, power input, and so on. Genomics encompasses fast automated sequencing, functional sequence analysis, and genome annotation.

The basis of biosystems engineering resides in the molecular structure of biological processes and structures (Figure 1.1). Problems and challenges of biosystems engineering have been addressed recently by the Stephanopoulos group at M.I.T. Their review [102] includes goals such as the improvement of cellular properties, the intelligent design of biochemical pathways and the eventual design of new phenotypes and the engineering of microbial cell factories to produce fuels, chemicals, and pharmaceuticals.

Biosystems engineering stands in contrast to a completely empirical approach for many microorganisms in which rounds of mutagenesis and genome shuffling are combined to optimize metabolic properties of the organism with respect to biotechnological parameters. This can be directed to a single enzyme substrate specificity, or simultaneously to several genes from related organisms, or entire chromosomes. These approaches are based on the "sexual PCR" (DNA shuffling) method pioneered by Willem "Pim" Stemmer (*1954–2013) founder of Maxygen [103].

Information flux		**Mass flux**	
Storage	**Message**	**Product**	**Metabolite A**
DNA	*mRNA*	*Protein*	
(Gen)		*(Enzyme)*	
ATG·CTC …	AUG·CUC …	Met·Leu …	Metabolite A
5000	4000	4000	500–2000
Genome	Transcriptome	Proteome	Metabolome

Figure 1.1 Basic molecular structure of biological production processes with microorganisms.

Synthetic biology is proceeding toward low cost and high volume with expectations for the market to reach nearly $12 billion in 2018, notably using synthetic DNA in organism design [104].

A number of examples of biosystems engineering applied to industrial biotechnology (or "industrial systems biology") have been summarized by Papini *et al.* [105] and Wittmann [106], including examples for commercially highly important products such as ethanol, butanol, polyhydroxyalkanoates, amino acids, polyketides, and antibodies.

The information stored in DNA is transcribed into messenger molecules (mRNA) that in turn encode the synthesis of proteins on the ribosomes. Most of the proteins produced act as catalysts for the reactions in the metabolic network. The entirety of the DNA information is known as the *genome*; for microorganisms, it typically contains from 1000 up to 5000 identified genes. The proteome comprises about 4000 proteins and roughly 2000 metabolites can be identified as comprising the metabolome [100].

Metabolomics is an extension of proteomics and describes the analysis of the catalytic activities that make up the metabolic activities of a cell. Metabolomics is at the core of biosystems engineering and its purpose is to analyze the metabolome (all metabolites, roughly in the range of 500–2000) and their concentrations in the cell under given physiological conditions, as well as the dynamic response to changing environmental conditions (notably in a bioreactor).

Amino acids, citric acid, lactic acid, propanediol, penicillin G, synthetic drug intermediates, and therapeutic proteins are among the industrially relevant products of fermentation and cell culture that have been targets for metabolic engineering. Some of this work has been adopted by industry (see [72], section 16.4.1). The major aim was to optimize the yields of industrial products, which was efficiently realized with *Corynebacterium glutamicum* for lysine and tryptophan, and at the Dupont company for 1,3-propane diol production [107–109].

Bioinformatics aims to effectively deal with data pools from the omics to provide useful interpretations and models. Bioinformatics provides informational techniques specifically enabling access to and interpretation of large amounts of data generated in different fields of biosystems engineering. It is used to develop tools for annotating large genome sequences, analyzing data produced by sequencing machines within a short timescale, interpreting data obtained from proteomics (two-dimensional gels and MS analysis), calculating mass balances in MFA, and determining the kinetics and dynamics from the analysis of metabolic pathways. It comprises understanding and modeling of genomes, proteomes, cell metabolism, and whole-cell biotransformations, and even incorporates the design of cell factories, in addition to modeling cells for new processes and for the formation of new products [110].

The scientific status of biotechnology can actually be recognized as that of a scientific discipline on its own, this being due to the significant progress in the scientific basis, molecular biology, as well as biochemical and biosystems engineering (see Section 1.4.5.1).

1.4.4 Products

The range of products has increased significantly since the 1970s, in the fields of food, feed, industrial commodities, specialties of the pharmaceutical industry, and agriculture. Moreover, services in environmental protection, notably wastewater and exhaust gas treatment, have developed into a large industry (for an extended overview, see Buchholz and Collins [72], chapter 16). Recombinant technology has led to considerable improvements in processing, product purity, reduced costs and prices, and to an extension of the range of products available, notably of pharmaceuticals and in-plant breeding.

1.4.4.1 Food, Feed, Industrial Commodities, and Specialties

For hundreds of years, industrial biotechnology has been a large, traditional technology concerned with producing beer, wine, and bread (see preceding sections), but has nevertheless been a dynamic field which now makes use of the most advanced tools such as rDNA technologies and systems biotechnology. Industrial biotechnology, often termed *white* biotechnology, comprises a broad spectrum of products and processes:

- Fuel, for example, ethanol and biodiesel, and energy, for example, biogas
- Commodities: organic acids, acrylamide, detergents, biopolymers, for example, biodegradable PLA and polyhydroxybutyrate (PHB), and enzymes
- Food and feed ingredients, textile and paper, for example, amino acids, starch derivatives, and sweeteners
- Fine and speciality chemicals, for example, antibiotics, chiral intermediates for pharmaceuticals, agrochemicals, sugars, and derivatives such as sorbitol, specialty enzymes, vitamins, dyes, fragrances, cosmetics, and polysaccharides.

It also aims at the utilization of renewable resources with the challenge of replacing oil-based chemicals and fuels. Environmental processes play a major role in industrial biotechnology, enabling clean and sustainable production and, furthermore, providing a safe environment most notably in crowded areas, big cities, and industrial areas. These technologies represent big business.

Table 1.2 shows a small selection of products; however, many more are actually marketed. For conventional food products, such as beer and wine (about 195 million tons per annum), cheese (about 19 million tons per annum), bakers yeast, vinegar, and others (see [73], chapter 1 table 1.1).

Although the production of biofuels is of considerable economic relevance, it is also characterized by heated political, ecological, and social debates. Recombinant technologies and second-generation fuels are expected to contribute to a reduction in the dependence on fossil resources and to significantly reduce greenhouse gas emissions ([73], section 12.2). A calculation of emissions yielded 94 for gasoline, 77 for currently available bioethanol, and 11 for cellulosic-based ethanol

Table 1.2 Selected products made by fermentation (worldwide, 2003–2005) [72, 96, 111–114].

Product/process	Production ($t\,a^{-1}$)	Price (€ kg^{-1})	Market value (million€)	Company
Ethanol	37 500 000	0.4	15 000	Diverse
Starch products	>10 000 000	—	—	Diverse
Glucose	40 000 000	—	—	Diverse
L-Glutamate	1 500 000	1.20	1800	Ajinomoto, Tanabe Seiyaku (Japan)
L-Lysine	850 000	2.00	1400	Evonic (D)
Citric acid	1 100 000	0.80	880	Diverse
Enzymes	—	—	1830	Novozymes (DK), Genencor (USA)[a)]
HFCS	8 000 000	0.80	6400	ADS, A.E. Staley, Cargill, CPC (USA)
Isomalt	>100 000	—	—	Suedzucker, Cerestar (D)
PLA	140 000	2.52	315	Evonic (D)
Xanthan	40 000	8.40	336	
Penicillins	45 000	—	Total antibiotics: 19 000	DSM(NL), Bayer (D), Kaneka (Japan), a.o.
Cephalosporins	30 000	—	—	
Riboflavin (B2)	30 000	—	—	BASF(D), DSM (NL)
Vitamin C	80 000	8	640	Roche (CH), BASF (D), Takeda (Japan)

Note the inverse relationship between production volume and price.
ADS, Archer Daniels Midland; CPC, CPC International; HFCS, high fructose corn syrup and/or glucose/fructose syrup sweetner; Isomalt, hydrogenated isomaltulose sweetner; and PLA, polylactic acid.
a) Taken over by DuPont (USA).

(measured in kilogram of CO_2 equivalent per MJ of fuel production and burning) [115]. Ethanol is both currently and traditionally one of the most important biotechnology products to emerge since the nineteenth century, although its production has been variable. A total of 46 million m^3 of ethanol was produced worldwide in 2006 and currently 47 and 28 million m^3 are being produced in Brazil and the United States, respectively. To improve the existing technology, the challenges include the optimization of *Saccharomyces cerevisiae* strains using recombinant technologies and obtaining higher osmotic and alcohol tolerance and sufficiently high fermentation rates (yeast being very sensitive to alcohol concentration of $100\,g\,l^{-1}$). In order to identify and use new raw materials, the concept of integrated biorefineries producing commodities from renewable resources has been envisioned. This concept involves processing various feedstocks derived from a range of different biomolecules, including lingocellulosics such as wood straw and whole crops (e.g., maize), into a variety of useful products. The manufacture of such products, which include fuel, energy and materials, building blocks for

chemical synthesis, and chemicals, would make use of the entire biomass including by-products [116, 117]. The concept is still in the phase of pilot studies, and the future development of economic conditions will control its realization.

Methanogenesis was detected by Volta in 1776 but it was over 100 years later that this biological process was first exploited technologically. Actually, biogas could, in principle, provide the solution to energy production from any complex biomass, including the huge reservoir of available residues and wastes, estimated at a potential of some 100 million tons per annum (oil equivalent) [118]. Recently, microbial fuel cells (MFCs) have been developed as a most promising concept that makes use of microbial catabolic activity to directly generate electricity from the degradation of organic matter providing access to cheap and environmentally friendly energy sources [119].

Among traditional bulk products glucose is used for fermentation to ethanol, amino acids, and organic acids, with main applications in the food industries, and so on, and as a sweetener, manufactured on a scale of some 40 million tons per annum. Furthermore, it is isomerized enzymatically to give a glucose/fructose syrup (high fructose corn syrup, HFCS) in a process that had been developed and scaled up to very large dimensions during the 1980s ([73], section 8.4). Similarly, the enzymatic isomerization of sucrose to give Isomalt has been established on a large scale. Amino acids, as additives for food, feed, and organic acids, with main applications in the food industries, have been produced in large quantities for many decades, mainly by fermentation but also by enzymatic processes, in the range of over 1 million tons per annum. The success in improving yields by metabolic engineering has been mentioned in the preceding text. Antibiotics, such as penicillin and cephalosporin and their derivatives (e.g., aminopenicillanic and aminocephalosporanic acids), are considered as bulk products in general, as the quantities produced are high and the prices are rather low, in contrast to most pharmaceuticals. New products entering the market during recent years include lactate and 1,3-propanediol as building blocks for polymer synthesis. Thus, the manufacture of PLA, which is based on lactic acid as the monomer, has been established on a large scale. The enzymatic production of 100 000 tons per annum of acrylamide from acrylonitrile was an important economic breakthrough.

For enzymes, a boost occurred from the 1980s on, with recombinant technologies extending the range of enzymes available. The new techniques considerably improved yields, lowered prices, and markedly extended applications, with modified and/or improved selectivity, effectiveness, and stability (both at elevated temperatures and pH). The current market for technical enzymes is about €2.5 billion, due to a significantly extended range of applications. The most important application has been in detergents (34%) and in starch processing to make sweeteners (some 30%), and, later, biofuels. During recent decades, their applications have extended steadily, owing to their excellent stereoselectivity, into the manufacture of fine chemicals (see the following) ([73], p. 15; for details, see chapters 1, 7, and 8).

A most important innovation came from academia, that is, the immobilization of expensive enzymes to give biocatalysts of high stability, high half-life, and low

cost, which enabled continuous processing, a breakthrough for large processes. Remarkably, one of the pioneers, Georg Manecke of the Berlin University, found no interest when he offered a patent to the Bayer Company in the mid 1950s. Some 15 years later, another pioneer, Günter Schmidt-Kastner, developed a process for penicillin hydrolysis, to be used in the synthesis of penicillin derivatives at Bayer. This process was successfully scaled up and went into production on a large scale around 1980; it was also developed at Beecham, working together with Malcolm Lilly of Imperial College London, and by other companies [120]. The first process with immobilized enzymes went into production in Japan in 1969 [121]. The largest enzymatic process is glucose isomerization to give a fructose/glucose syrup with production of some 8 million tons per annum.

Cloning of the first industrial enzymes, penicillin amidase and α-galactosidase, both in *Escherichia coli*, was achieved respectively by Wagner, Mayer, and Collins, Braunschweig and by Bückel, Boehringer Mannheim during the late 1970s [122]. Large-scale production of α-galactosidase was carried out at Boehringer Mannheim and for amylases at Novo in the early 1980s. The creation of modified or new activities via generation and design of mutants, directed evolution, and rational design was very successful in the production of chiral chemicals as synthons for pharmaceuticals and agrochemicals. Enzyme discovery, screening, and selection methods have undergone much progress (Bornscheuer, in [73], chapters 3 and 5) (for the history of enzyme technology, see [120, 123]). Hubert Mayer, in the late 1970s, had tried to clone penicillin acylase (PA) without success, an idea that came from Fritz Wagner; the gene of PA at first was not found in the clones. The key for success was the cosmid technique introduced by John Collins, with which the whole *E. coli* genome could be obtained within 200–400 clones ([72], p. 123; [123]). The cosmid patent was filed immediately in 1978, and was the third rDNA patent to be approved worldwide, an entrepreneurial act received negatively by the academic community. Such expression cloning is still used for "empirical gene mining" using DNA from poorly characterized collections (mixtures) of commensal species from various environments even where individual cultivation of isolated species is not possible.

Around 1980, genetic engineering of amylase production was a test case at Novo for food enzymes, going through the approval process, and finally being marketed in 1984.

Fine chemicals made using biotechnology comprise a wide range of products. The more recent innovations include vitamin B12 and the expanding field of chiral organic building blocks for pharmaceuticals (see [72], section 16.4.3; [73], chapter 4). Since the mid-1970s, a new approach and impetus to systematic studies aimed at application of biocatalysis and biotransformation in organic synthesis was undertaken. It included semisynthetic penicillins, esters, and glycerides [124], as well as optically active compounds [125]. It was recognized that biocatalysts could be used for synthesis in organic chemistry, notably for pharmaceuticals and agrochemicals incorporating chirality, that is, the production of pure compounds consisting of one isomer only as the key requirement. An important economic breakthrough was the enzymatic synthesis of chiral

aspartame, an intense sweetener, the annual sales of which are in the region of $850 million [126].

Environmental processes play a major role in industrial biotechnology, both enabling clean and sustainable production in industry and providing a sound environment. They comprise wastewater and exhaust air purification as well as soil remediation, and they represent big business ([127, 128], pp. 309–321). Thus, environmental biotechnology is of major importance, considering that all cities worldwide and every town in western countries is obliged to undertake biological wastewater treatment, and that most industries including not only food, pharma, and chemistry but also car factories have a requirement for biological wastewater treatment and in many cases for biological exhaust air treatment systems as well. In 1988, there were about 27 400 aerobic municipal wastewater treatment plants in the European Union. The cost of wastewater treatment worldwide has been estimated to be $30 billion in 1980 and $68 billion in 1990, with an annual growth rate of 9%.

1.4.4.2 Biopharmaceuticals

The setting up of new biotechnology companies (Section 1.4.2) was the key event for the development of biopharmaceuticals based on recombinant technologies. Antibiotic manufacture had been established before this (Section 1.3). Antibiotics are still the most important strategy against bacterial infections. The world production of antibiotics is estimated at over 60 000 tons per annum, valued at more than $30 billion per year [128, 129]. For the synthesis of semisynthetic penicillins and cephalosporins, the fermentation products penicillin G and cephalosporin, respectively, are hydrolyzed by immobilized enzymes to yield the acid form with the intact β-lactam ring as the active principal: 6-APA, 7-ACA, or 7-aminodeacetoxy cephalosporanic acid (7-ADCA).

Protein-based drugs constitute about a quarter of new approvals with a majority being glycoproteins. A total of 77 recombinant biopharmaceutical products gained approval and were available on the market in 2000 and that number had increased to some 165 biopharmaceutical products by 2006, with a market size estimated at $35 billion in 2006, and reached over $60 billion by 2014 [130] ([72], sections 9.7.2 and 17.4.2). The average annual growth rate was 20% ([87], pp. 8–11; [130, 131] (for an overview see: [98]); see also www.centerwatch.com/patient/drugs). Approved recombinant proteins comprise a number of different biopharmaceuticals, that is, blood factors, thrombolytics, insulins, other hormones, hematopoietic growth factors, bone morphogenetic proteins, IFNs, and ILs, mABs and engineered antibodies, vaccines, therapeutic enzymes, and enzyme inhibitors [97]. Most of the drugs listed in Table 1.3 are blockbusters with sales of $1 billion per year or more. Most of the newly approved medicines have been developed to address indications which are the major causes of disease and mortality in the industrialized world, notably infections,

Table 1.3 Examples of rDNA products for medical use compiled and extended after [72, 87, 131, 133].

Active substance	Medical use	Gene first cloned	FDA: first approval[a)]	Market size ($\$ \times 10^9$ yr^{-1})
First generation				
Human insulin	Replacement therapy diabetes type 1/2	1978	1982	12.15[b)]
Growth hormone (huGH)	Replacement therapy dwarfism	1979	1985	1.02[c)]
Factor VII/VIIa	Replacement therapy, hemophilia (Alexander's disease)	1987	1999	0.30[d)]
Factor VIII	Replacement therapy, hemophilia A	1988	1993[e)]	2.40, 3.00[d)]
Factor IX	Replacement therapy, hemophilia B	1983	1996[f)]	0.30|[d)]
Enzymes[g)]				1.30[g)]
β-Glucocerebrosidase	Gaucher's disease	1985	1994	0.20[h)]
α-Galactosidase	First drug for treating Fabry disease	1985	2003	0.20[i)]
Iduronate-2-sulfatase	Replacement therapy mucopolysaccharidosis type II	1990	2006	0.20[j)]
Human DNAase	Lung clearance, cystic fibrosis (mucoviscidosis)	1990	1993	0.30[k)]
Second generation				
Human interferon-α1	Antiviral, limited antitumor (hairy cell leukemia)	1979	1984	0.60[l)]
Human interferon-β	Antiviral (HCV), multiple sclerosis (MS)	1979	2002[a)]	3.80[m)]
Interleukin-2	Metastatic renal carcinoma	1983	1992[n)]	0.08[o)]
Tissue plasminogen activator (tPA)	Dissolves blood clots (e.g., after heart attack)	1983	1996[a)]	0.85[p)]
GM-CSF	Shortens neutropenia after chemotherapy	1985	1991[a)]	0.08[q)]
G-CSF	Use in blood stem-cell allografts	1986	1991[a)]	4.50[r)]
Erythropoeitin (EPO)	Replacement therapy (EPO lost during hemodialysis)	1985	1996	3.50[s)]
Recombinant vaccines				1.43
Hepatitis B vaccine[t)] (HBS antigen)	First anticancer vaccine (80% of hepatic cancer → viral origin)	1982	1986	0.20[u)]
Papilloma virus vaccine (HPV)	Anti-cervical cancer vaccine	1980	2006	0.60[v)]

Monoclonal antibodies (mAbs)				24.60
Anti-TNFα	Autoimmune disease (plaque psoriasis, rheumatoid arthritis)	1991 (Enbrel)	2002[a)] (Humira)	11.00[w)]
Anti-HER-2/neu	Metastatic breast cancer, gastric cancer ($HER2^+$)	1984	1998	1.90[x)]
Anti-CD20	Non-Hodgkin lymphoma, B-cell leukemias	1994	1997	3.50[y)]

a) The first FDA approval may be for other indication(s).
b) Insulins (Lantus, Novolog, Humalog, Levemir, Humilin) plus GLP-1 receptor agonists (Victoza, Byetta, Novolin).
c) Human growth hormone (nutropin, genotropin, humatrope, norditropin).
d) Blood factors total about $3500 million, Factor VII (Novoseven,); Factor VIII (Advate, Refacto AF/Xyntha); and Factor IX (Benefix).
e) Human albumin-free first in 2003.
f) 2013 Version with long serum-half-life approved here.
g) Therapeutic enzymes, total US sales (2012).
h) Cerezyme + Vpriv.
i) Fabrazyme + Replagal.
j) Elaprase.
k) Pulmozyme, the first inhaled protein therapeutic. Interferons/cytokines.
l) Pegasys.
m) Avonex, Rebif, Betaseron.
n) Delay in approval caused failure of Cetus (bought by Chiron); US market data updated to 2012.
o) Proleukin, 2010.
p) Alteplase, worldwide.
q) Granulocyte-macrophage-colony-stimulating factor, leukine.
r) Granulocyte-colony-stimulating factor, Neulasta, neupogen.
s) Epogen, Aranesp, Procrit.
t) N.B.: PCR-based testing of blood and blood-derived products for viral contamination testing from 1986 $>10^6$-fold more sensitive, also for previously unknown viruses (HBV, HIV, HCV, HDV, West Nile Virus, etc.).
u) Pediarix, Recombivax HB.
v) Gardasil; monoclonal antibodies (US sales).
w) Humira (first fully human mAb) + Remicade + Enbrel (fusion protein), *world-wide sales $23 billion*.
x) Herceptin.
y) Rituxan, chimeric murine-human mAb.

with hepatitis B and C as the most targeted indications, diabetes, hemophilia, myocardial infarction, and various cancers [132].

Following a number of scientific breakthroughs, about 10 new products were approved in 2014, some with blockbuster status, comprising antibodies, peptides, and enzymes, for example, for treating cancer [134]. However, R&D costs to bring a single drug to market is approaching $3 billion, a 145% increase over estimated costs in 2003 [135].

1.4.4.3 Plant Products, Seeds

The promise of transgenic plants, or green biotechnology, is to create crops with higher yields, which can grow on less fertile land, to feed a growing, hungry world population. Headlines such as "Genes to gasoline" provoke expectations for second-generation biofuels. Crops should be resistant to pests and need fewer chemicals, notably insecticides, fungicides, herbicides, and fertilizers. The use of genomic analysis and quantitative trait linkage has not only accelerated plant breeding programs, but allows definition of new cultivars that need standards for patent protection, although no rDNA methodology was directly applied, that is, the plant concerned is not a genetically modified organism (GMO). The majority of agricultural scientists are convinced that crops of high yield, high quality, low cost, and low environmental impact can be delivered by the exploitation of the techniques for plant biotechnology, in particular by involving molecular breeding strategies ([136], Foreword).

The most successful, in terms of application, has been herbicide resistance, where reduced utilization of herbicides has been claimed. Resistance against diseases (pests) has been one of the priorities in designing genetically modified (GM) plants. This also implies a decrease in the dependence on chemical pesticides. Genetic modification has been shown to provide an additional and less time-consuming tool, and to be a useful means of creating novel resistance, compared to classic breeding.

Heated controversies and media attention have affected the discussion on GM crops. Few novel technologies have provoked so much opposition, and consequently governmental control, as has plant biotechnology. It has been subjected to major political, economic, social, ethical, and environmental scrutiny, much more than other fields of biotechnology. The discussion of risks in general goes back to the formulation of rDNA guidelines in Asilomar in 1975. The critical public perception and resistance of consumers throughout the world has slowed down, and even prevented applications, notably in Europe.

Applications include food, feed, and renewable resources, with over 15 crops modified for 1 or more of some 47 phenotypic traits commercialized since 2000, most with attributes related to input and yield performance. In 2006, 102 million ha were cultivated with GM crops, and in 2007, there were 114 million ha in 22 countries. Seven countries dominate the global market, namely, the United States, Argentina, Brazil, Canada, India, China, and South Africa. Among the products are soy bean, maize, cotton, rape, cucurbit (squash), tomato, papaya,

alfalfa, poplar, petunias, and paprika. The growth of the area cultivated since 1996 is obvious. Application was mainly with the combined traits of insecticide resistance and herbicide tolerance. Four crops account for nearly 100% of the commercially grown transgenic crops, namely, soy bean (64%), cotton (43%), maize (24%), and rape (20%).

1.4.5 Further Aspects

The focus of this section has been on the scientific and engineering foundation, background, and development of biotechnology. Nevertheless, some aspects of political and economic conditions, and of sociology of science should be mentioned. Thus, a considerable number of approaches tend to investigate, and find general interpretations and explanations for this unique technological development. They are not discussed here, but mentioned only as suggestions for further reading.

1.4.5.1 Scientific Status

Since the 1990s, the status of biotechnology can be recognized as that of a scientific discipline of its own. Advanced groups in applied biology, biotechnology, and biochemical engineering tended to cooperate and integrate their efforts, with work oriented to the molecular level and with empirical data being interpreted in terms of coherent theories [137, 138]. The basis was molecular genetics, which introduced a new paradigm based on knowledge of the molecular biology involved, and basic engineering concepts, which was oriented at molecular processes, and the integration of these scientific approaches [139].

1.4.5.2 Political, Institutional, and Socioeconomic Conditions

A general overview on the history of biotechnology, including these topics, was written by Bud [2]. He discussed the role of governmental politics, under the heading "Resistance to New Technology – Regulators of Biotechnology" [74]. A critical comparison of governmental politics in Germany, termed a *corporate approach*, and the United States, with the OTA study, was presented by Jasanoff [77]. Yi [140] discussed how academic institutions, government agencies, and the nascent biotechnology industry contested the legal ownership of rDNA technology in the name of the public interest. It meant the reconfiguration of the ownership of public knowledge in late twentieth century American capitalism. (See Section 1.5 for details of the earlier discussion.)

1.4.5.3 Economic Studies

A range of socioeconomic studies have been published in recent years, including the roles of academic, notably university, and company research, networks, and

company strategies, both of big pharma and new biotechnology firms However, many studies suffer from a lack of interdisciplinarity in research teams, where it seems that no scientists or only few scientists involved in biotechnology participated.

The group of Zucker and Darby [141–143] found a positive impact of research universities on nearby firms related to a positive impact on the firms' research productivity, increasing the number of biotechnology patents, more products in development, and more products on the market, as of 1989–1990. Arts and Veugelers [144] and Arts *et al*. [145] discussed indicators derived from patent documents to capture the nature and impact of technological inventions, and to compare and validate these indicators within the field of biotechnology. However, the studies suffer from a lack of interdisciplinarity, which may be considered necessary for such work.

1.4.5.4
Science Studies

What drives science and what motives make up its dynamics? Several periods in the history of biotechnology may be interpreted in terms of a concept developed by Rheinberger [146] ([72], part I), who analyzed the early phases of molecular biology, that is, Zameznik's approach to identify the steps in the translation of DNA into protein, which turned out to be a highly complex sequence of reactions. In this concept, epistemic things may be characterized as a field of certain, yet unidentified objects, experimental techniques, and implicit knowledge, that is, a field of problems of phenomena, reactions, and structures which were not understood unequivocally, nor explained. Adequate methods of working on the topic to give unambiguous results were not available. "To put it as a paradox: epistemic things represent that what remains unknown. They may be characterized by a list of activities and properties." Epistemic things may be considered as those questions, problems, and stimuli that provide the dynamics of research in early phases. No ideas of technical application in molecular biology appeared as long as the field remained unclear and undisclosed; these are, epistemic things in Rheinberger's notation. It was only when structures of DNA and protein and enzymatic reactions appeared with consistent character and logical order that rational construction seemed possible, and pioneers began to conceive concepts for application, that is, cloning concepts for making drugs and other proteins.

1.5
How Pioneering Developments Led to Genetic Engineering

1.5.1
Preamble: People and Principles in Developing Enabling Technologies

Bartneck and Rauterberg [147] have argued that emphasis on discoveries rather than on inventions has moved the Nobel Prize away from its original intention of rewarding the greatest contributions to society. It should be clear from the

following descriptions of events in molecular biology, molecular genetics, and cell biology that both novel conceptual insights and technical innovations have led to "enabling technologies" which have, to say the least, transformed biotechnology and modern medicine.

Berg and Metz [148] have presented gene technology as having arisen mainly as a consequence of using established tools in new combinations. This does little to explain why. In the following discussion, we differentiate between developments that have occurred by gradual incremental steps, perhaps involving less intellectual input, and those which arose by more radical conceptual changes, even when they mimicked mechanisms used in nature. Contrary to Berg and Metz, we argue that at the inception of gene cloning, it was not an accepted notion that, in the face of enormous ignorance of the basic differences in gene structure and function in different species, cross-species gene cloning, for example, selecting for expression was a likely possibility. Another category of developments involves surprise findings in the early years of gene cloning that caused redirection and creation of whole new fields of investigation. This led to novel enabling technologies beginning in 1972–1973. A special aspect unusual for the biological sciences at that time was that academic discoveries were immediately translated into biotechnology inventions that were taken up rapidly in industrial settings, producing enormous innovation, often with novel medical applications. This involved, in many cases, academics who became entrepreneurs, a fact that was vehemently condemned by many of their peers as *infra dig* and/or, especially with the patenting of sequences and products (necessary for the due diligence process required by potential investors in start-ups), as incompatible with scientific rigor and/or academic freedom.

1.5.2 Academic Freedom and Patenting: Hindrance to Science or Lever for Innovation?

This latter conflict area, although not suffered by Louis Pasteur (a great model to follow?) has occurred often before. Are there parallels to be seen with the conflict between Humphrey Davy and George Stephenson about the invention and patenting (1815) of the miners' lamp that protected miners from igniting methane explosions? Davy refused to patent and received the Rumford medal for this invention in 1816. However, his lamp was dull and deteriorated in damp mines. Stephenson's patented design invented independently, was bright, worked well, and was preferentially manufactured and used. Perhaps a similarity exists with Paul Berg being awarded a Nobel Prize for being the first to discuss cloning principles and doing *in vitro* DNA joining: In contrast, Cohen and Boyer invented and demonstrated a reproducible DNA cloning method, which they patented, opening up a multibillion dollar industry. James Watson (already a Nobel Laureate), the coordinator of the Human Genome Program at the NIH, was incensed by the National Institutes of Health patenting hundreds of Craig Venter's EST DNA sequences. Watson often referred to Venter as "Hitler," later saying in public that "virtually any monkey" could do that sort of work (EST mapping; [149]). Venter moved genomic

analysis into an industrial setting on a grand scale and now uses a fund derived from this to donate more financial support for international health programs than the World Health Organization itself.

1.5.3 Conceptual Leaps and Jumps

Enzymes that had been discovered to be part of cellular defense mechanisms to recognize and destroy foreign DNA were reconceived as tools to ensure its uptake into viral and plasmid vectors, thus allowing its stable perpetuation across species barriers: this was the beginning of the ability to manipulate DNA at will. This was true both during 1972–1973, with the use of restriction endonucleases (R.endos) and, more recently, since 2010, with the use of CRISPR/Cas9 (a type of bacterial adaptive immune system recognizing specific fragments of foreign DNA). This latter system has recently been developed into the most specific method for genetic engineering and site-specific mutagenesis in a wide range of prokaryotic and eukaryotic hosts [150] with or without plasmid or viral vectors.

In the absence of knowledge of gene structure and function in higher organisms, rDNA technology was used as a basis for a new industry to produce pharmaceutical proteins at a time when the pharmaceutical industry thought this would be impossible. In retrospect, this would seem to require that we recognize this step as a conceptual leap, although the individual components needed for constructing rDNA molecules are known. A gene is not just a piece of DNA. It is a *particular* piece of DNA that shows stable inheritance and, when in the *correct chromosomal location* in a *specifically developed cell* containing the *required regulatory molecules and receptors*, can be expressed in a living organism, preferably in a controlled fashion. Its protein product may require specific modification before it can be functional and may require extensive purification before it is suitable for use as a pharmaceutical product.

At several steps on the way to acquiring the human genome sequence and finally putting this information to biotechnological use, a conceptual breakthrough involving several developments were needed. The following are some examples:

- Botstein's analysis of how high-resolution genome mapping could be achieved and used for human genetics and gene isolation [151];
- rapid and cheaper analysis made possible initially by better cloning systems for efficient isolation of larger fragments;
- better computer systems and algorithms for sequence analysis that allowed shotgun whole genome analysis;
- the EST concept which allowed more rapid access to the protein-coding regions of genes (not forgetting Sydney Brenner's novel contribution sequencing *Fugu* the Puffer-fish, whose genes essentially lacked introns → 10-fold less sequencing to get to the gene sequences);
- novel concepts for faster parallel sequencing to remove the requirement for previous cloning of large DNA regions;

- the brilliant idea of the PCR, which was completely unexpected [152, 153] and hardly believed before it became generally established with the introduction of a heat-stable DNA polymerase (Taq1), which became generally available through cloning its gene in 1989 [154]. This method allowed, among other things, direct cloning and sequencing of any region bracketed by known sequences (this is known as simplified gene manipulation), *in vitro* affinity selection of gene and gene product and thus isolation of completely novel ligands with specific affinities or functions;
- affinity selection and affinity maturation that involved novel brilliant concepts in the area of phage display and ligand libraries, mimicking natural processes found in antibody maturation, or seen over long periods of time in evolution. For extensive reviews, particularly for the beginnings of gene cloning, phage display, applications in biotechnology, and recombinant antibody development, see [72, 155–157].

1.5.4 Surprise Discoveries Initiate Novel Areas/Methods of Research

A number of discoveries were completely unexpected:

- *Gene splicing*: Many coding regions in eukaryotic genes and particularly their viruses are often not continuous, but interrupted by intervening sequences that have to be removed by a *splicing* process at the mRNA level before translation takes place. Gene splicing was discovered independently by Sharp [158] and Roberts [159] in 1977.
- *Nonuniversality of the genetic code*: One of the assumptions implicit in doing gene cloning is that the genetic code is universal. Luckily for those who had invested in this new technology, this was generally true. Exceptions existed, however, for example, as discovered in bovine mitochondria cytochrome oxidase II where a stop codon translated to tryptophan insertion (see, e.g., a retrospective by Sanger [160]): Later more exceptions were found: for example, four base-pair suppressor tRNAs were found to be commonplace in yeast; and in bacteria, the formation of seleno-enzymes.
- *RNA enzymes*: Studying splicing also led to the discovery of *RNA-enzymes*, interpreted by many as an insight as to what might have been used as the biochemical basis of life before the evolution of complex protein synthesis machinery. This finding contributed to the development of SELEX technology and "synthetic biology."
- *Gene transfer to plants*: The studies of tumors in plants by Jeff Schell and Marc van Montagu led to using the same system that one finds in nature for transfer of genes from bacteria (*Agrobacterium tumefaciens*) to plants, the beginning of genetic engineering in plants.
- Gene copy number variation (GCNV): the discovery that GCNV between individuals is a major factor in human genetics was unsuspected. This complicated

the genome-sequencing program that was based on mixed DNA from many different individuals. It now influences the way human geneticists screen patients; for instance, DNA chip arrays can be used for this purpose.

- *Missing heredity*: (Low clinical relevance of genetic predisposition): The modest results in finding genes strongly affecting widespread syndromes such as metabolic syndrome and susceptibility to cancer was also not expected. Continued huge funding in this area meets significant criticism [161], although many are still looking through rose-tinted glasses at the $2 billion increase in the US Congress budget for Vice President Biden's "moonshot" to defeat cancer.
- *Epigenetics*: Changes in environment cause changes in gene expression via, for example, alteration of methylation at the DNA level and through modification of histone and other chromatin proteins. These so-called *epigenetic modifications* now represent an important new area of research and diagnostics with high relevance in health care.
- *Unknown microbial flora*: Shotgun sequencing of DNA from entire microbial communities revealed a vast unrecognized microbial variation: microbiologists, until the 1980s, had largely confined their studies to microorganisms which could be cultivated as pure colonies in the laboratory. The astounding finding was that perhaps more than 95% of the world's microbes had not yet been isolated in pure culture! Cell growth is dependent on commensal and symbiotic interactions or, for some, cell growth stops at very low cell density ("quorum sensing"). Amazingly, the *most abundant single species in the world*, *Pelagibacter ubique*, was first discovered by assembly of the complete genome sequence from sequences found in nearly all ocean samples. The huge potential represented by this wealth of microbial diversity and specialization is a boon both for basic research and for the biotechnological industry concerned with secondary metabolites. It represents a rich area of specialization in terms of gene mining for many years to come. The Earth Microbiome Project (EMP) was launched in August 2010, with the ambitious aim of constructing a global catalog of the uncultured microbial diversity of this planet. The primary vision of the EMB, to process the microbial diversity and functional potential from approximately 200 000 environmental samples, marks it as an undertaking so massive that it was at first considered to be pure folly (as late as 2012, Jonathan Eisen was quoted in Nature as saying, "Knight and Gilbert literally talk about sampling the entire planet. It is ludicrous and not feasible – yet they are doing it" [162]).

1.5.5 Methodology Without Which Gene Technology Would Not Have Been Possible

Before dealing with the specific steps that made gene technology possible as a direct precursor to genetic engineering, we should consider the methodological repertoire that had mostly been developed in the beginning and middle of the twentieth century. The standard laboratory methods are often taken for granted, but a quick review of the most important is included here: their effects on discoveries are listed in Table 1.4.

Table 1.4 Laboratory procedures on which the growth of recombinant DNA was also dependent.

Chromatography (see body text 1930–1960s)
- Paper/starch → first analysis of nucleotides and chemical synthesis products; inherited changes in proteins; + antisera → antigen recognition (Ouchterlony method)
- Size exclusion polymers and zeolites → pure enzymes
- HPLC and gas chromatography → combination of the above were the basic tools of biochemistry combined with structural analysis: sequencing of first simple RNA viruses; peptide sequencing

Electrophoresis: agarose/acrylamide gels → protein separations; isolation of specific DNA restriction fragments; first DNA sequencing; analysis of transcription products (Northern blotting); or proteins (+Western blotting)
- Pulse-field electrophoresis → mapping very large (>30 kb DNA fragments e.g., NotI digest of YAC clones; jumping and linking libraries for genome contig building/mapping
- Isoelectrophoresis focusing (IF) → refinement of antigen recognition; protein charge shifts detected
- Combining 1D and IF → 2D protein gels → detect post-translational modification; proteomics tool

Centrifugation → preparation of cell pellets/antibody-antigen complexes
- Velocity sedimentation → fractionation of organelles; cellular components, RNA, DNA, membranes
- Ultracentrifugation (isopycnic) → DNA strand separation, first pure gene; preparation of plasmid/viral DNA

Electron microscopy (EMI) → DNA lengths
- EM + DNA, or DNA/RNA heteroduplex → discovery of introns, transposons (Tn), insertion elements (IS)
- *Scanning tunneling EM (1981)* → image individual atoms; confirmed DNA structure; future single molecule sequencing? Longer fragment sequences: faster sequencing?

Crystallography → recognition + separation of mirror symmetry enantiomers
- X-ray diffraction analysis → absolute molecular structure. First protein structures combined with powerful computing → protein design

Laser-activated fluorescent technologies
- Flow cytometry (cell sorting, 1972)[a] → first high-throughput screening; → with antibodies follow cellular differentiation; separate specific cell types
- Combined with tagged gene expression → follow cell differentiation *in vivo*
- Single molecule fluorescence (1992) → follow intracellular trafficking; next-generation single molecule sequencing? Novel methods for measuring intermolecular affinities, for example, in drug screening (Evotec)

a) First developed for electrical impedence (Fulwyler, 1965) → founding of Cetus.

1.5.5.1

Centrifugation: Preparation of Molecules of Different Sizes, Shapes, and/or Densities: Velocity Gradient Centrifugation

Friedrich Miescher is generally given credit for being the first (1869) to isolate a cell organelle by centrifugation [163]. The protease stripped extract of the "nucleus fraction" of human pus, which he called *nucleins*, was later renamed by Altmann as nucleic acid, once its chemical nature had been established.

Centrifugation (Table 1.4), initially applied in the dairy industry, and ultracentrifugation, now a routine method in all biology laboratories became a science through the work of Theodor Svedberg in Uppsala, and the unit used for the sedimentation constant (S = Svedberg = 10^{-13} s), reflecting the relative sizes of

molecules, is named after him. To study the size of proteins and RNAs, very high g-forces had to be generated, for example, already in the 1930s, 900 000-fold greater than the standard gravitational field. The older authors of this chapter often experienced rotors exploding in their armored centrifuges during molecular (isopycnic) separations in the 1960s and 1970s.

- *Velocity centrifugation*
  Behrens applied gradient centrifugation, where sugar gradients were used to prevent mixing of the tube contents during acceleration and braking. This was used for the fractionation of cells from blood or tissues and analysis and preparation of subcellular fractions, for example, nuclei, chloroplasts, ribosomes, mitochondria [163]. This early work was instrumental from the 1940s onward in the detection of molecular subunits of complex molecules such as the protein subunits of hemoglobin.
  Continuous preparative ultracentrifugation was used in the mid-1970s for isolating mitochondria, chloroplasts, ribosomes, or RNA species from cell lysates on a large scale.
- *Isopycnic density centrifugation*
  During high speed centrifugation (e.g., several days at >300 000*g*), even the salts of heavy metals, such as cesium, form a density gradient in which large molecules form a discrete sharp band at their specific density. This required that engineers develop a rotor that could sustain such huge forces [164]. With this method, it was shown that transfer of resistance properties between bacteria was sometimes associated with the acquisition of "satellite" or "episomal" DNA bands, differing in density from the chromosome [80]. Single strands of bacterial viruses could be separated in isopycnic gradients after melting the DNA and hybridizing it to synthetic poly (rI:rC). This was instrumental in isolating the first pure gene fragment in 1968, the pre-DNA cloning era (see below) [165].

In 1967, Vinograd was instrumental in analyzing plasmid DNA via ultracentrifugation, including the use of DNA-intercalating dyes such ethidium bromide [166, 167]. This allowed reliable detection and purification of plasmid or viral DNA on a large scale, even when it was of the same specific density (AT/GC ratio dependent) as the host chromosome. This was the main method in use from the early 1970s and in the mid-1980s. The fact that ethidium bromide-DNA complexes are highly fluorescent in UV light allowed sensitive detection of DNA in gels, facilitating restriction mapping and later the analysis of PCR reaction products (e.g., DNA fingerprinting developed by Alec Jeffries in 1985). It also introduced mutations in the DNA.

1.5.5.2 X-ray Crystallography: Understanding Molecular Structure at the Atomic Level

The discovery of diffraction of X-rays by crystals by Max Von Laue in 1912 was an indication that their atoms are arranged in ordered lattices. William Henry Bragg refined the X-ray spectrometer for use with X-rays of defined wavelengths. His son, William Lawrence Bragg, developed the mathematics to deduce the positions

of atoms within the crystal. This was initially calculated by hand for simple mineral salts. Extending this to proteins required faster computing, improvement of X-ray beam production (e.g., at the DESY synchrotron), and ever more sensitive high-speed data collection of the diffracted beams. The expense of the instrumentation allowed the establishment of the technology initially in relatively few centers. These became centers of excellence in molecular biology, attracting excellent scientists and resulting in rapid developments, for example, as seen from the fact that, between 1904 and 2012, 41 members of either the Cavendish Laboratory, or the Laboratory for Molecular Biology (LMB) at Cambridge University, UK, became Nobel Laureates (Fred Sanger twice; *recommended further reading*: (i) de Chadarevian [168], (ii) Brownlee [169]).

1.5.5.3

Chromatography with Solvent Motion or Electric Charge: Detection of Mutant Gene Products

Electrophoresis as a major analytical tool began with the work of Arne Tiselius. Tiselius, a student of Svedberg, developed moving boundary electrophoresis in 1937. This was based on Faraday's work on electrolysis in the eighteenth century and on studies of moving boundaries in electrolytes by "schlieren" optics, initially applied to protein analysis by Botho Schwerin in 1914 (Section 1.5.5.4). Oliver Smithies in the 1950s was the first to detect altered protein structure correlated with inherited diseases. In inherited thalassemia, those carrying the trait had altered globin conformation with lower affinity for oxygen that caused it to run at a different speed in zone electrophoresis through starch gels.

1.5.5.4

Protein Sequencing

Earlier developments in protein purification and chromatography had a lot to do with the development of protein sequencing by Fred Sanger starting in 1943 in Chibnall's group, Cambridge, UK. Insulin was already available in a pure form from the pharmaceutical industry. Sanger developed a sequencing protocol involving specific chemical modification of the exposed amino groups. After partial hydrolysis of insulin (e.g., with acid or proteases), he fractionated the products using 2D paper chromatography; in one direction by electrophoresis, and in the other by solvent. Ninhydrin stained the breakdown products for visualization on paper. These "fingerprints" could be interpreted in terms of a sequence, initially only near the amino termini of the peptides. The total sequence, including the position of disulfide bonds linking the two chains, was established by 1955. This work finally established that proteins had definite polypeptide chains of specific sequence and were not just amorphous. These data later, in combination with analysis of RNA and DNA structure, allowed the conceptualization that there is a flow of genetic information, each in the form of linear structures, from DNA to RNA, and from RNA to protein. Following more sequencing and the combined use of oligonucleotide synthesis and *in vitro* protein synthesis, the rules for this high fidelity information flow, the genetic code was solved.

1.5.5.5

Nucleic Acid Sequencing – the Prelude: Phage + Bacterial Genetics and Biochemistry, the Gene Concept

Short synthetic oligonucleotides and RNA viruses were the first, pure, single-stranded nucleic acids available for sequencing before gene cloning. Bacterial genetics in combination with biochemistry, to examine the composition of the protein gene products, allowed an initial analysis of genes and the discovery that some jointly regulated inducible gene clusters (operons) were sequestered on plasmids and bacteriophage which could be transferred (by conjugation, transformation, or transduction) between bacteria.

Proteins controlling gene transcription were isolated by binding to bacterial viruses (bacteriophage or "phage"), for example, for the lactose operon (Qβdlac) or for the phage λ (*lambda*) itself. In the late 1960s, the exact molecular and biochemical interaction controlling gene expression in the bacterial *lac* operon was established, 5 years before gene cloning was developed. A circular genetic map existed already for both the phage λ (*lambda*) and the bacterium *E. coli*. The first gene fragment was isolated in1969.

The first isolation of a pure gene fragment was carried out 4 years before gene cloning [165] at Harvard. This involved S1 single-strand nuclease digestion of DNA heteroduplexes (partial hybrids obtained after melting the dsDNA and reannealing at lower temperature) of transducing phages containing the lac operon in inverse orientation. This was hailed by the *New York Times* as the beginning of a revolution in planned genetics, similar to that portrayed in Aldous Huxley's novel "Brave New World," thus inaugurating self-criticism by the community involved in the research.

The Beginnings of Sequencing By the late 1960s, small RNAs, for example, tRNAs and RNA viruses, had been highly purified. Sequencing of the viruses by partial chemical degradation and chromatography of the products was initiated in Fiers and Weissman's groups in 1969. Sanger's group was instrumental in many of the developments in this early period. The isolation of a tRNA pure enough for sequencing required the newly developed DEAE cellulose ion exchange chromatography and further refinement of countercurrent chromatography. Once this had been achieved, Holley's group [170] was successful in sequencing the 77 base yeast alanine-tRNA, the first RNA molecule sequenced [171]. Studies on *in vitro* translation of linear single-stranded RNA or single-stranded DNA (ssDNA) viruses were instrumental in confirming the degeneracy and universality of the genetic code [172]. Sequences specifying the site of initiation of translation were determined by sequencing fragments of the RNA phages protected by ribosome binding from degradation by ribonuclease [173]. The first gene sequence, from an MS2 RNA virus, came from Walter Fiers' lab in Ghent, Belgium [174] in 1972, as did the first total genome sequence in 1976, again the MS2 RNA virus [175]. Sanger and Weissmann were pioneering oligonucleotide sequencing using highly radioactive ^{32}P-labeled RNA, complete or partial digestion with enzymes, and sedimentation and chromatographic separation of the cleavage products. The chromatographic separation methods were the same

as those used by those synthesizing oligonucleotides, for example, in Khorana's laboratory, and these synthetic fragments could be used as reference material [176]. Gilbert and Maxam in 1973, established the 23 bp lacO (operator) dsDNA sequence, separating chemical degradation fragments by chromatography [177]. Weissmann's work pioneered the chain-elongation sequencing technology using labeled nucleotides in reiterated viral RNA synthesis reactions lacking one or more of the four nucleotide-triphosphates.

The first RNA molecule to be sequenced was not a virus, but a tRNA. Two enzymes were required in these analyses: (i) bovine pancreatic ribonuclease, which cleaved after pyrimidines (C or U), became a classic system for scientific studies after Armour & Co., the hotdog Company, purified 1 kg of the enzyme and distributed it to scientists (source Wikipedia) and (ii) takadiastase ribonuclease T1, which cleaved 3′ to a guanosine (G). Each of these small fragments was further analyzed by partial digestion with snake venom diesterase from the 3′-ends. Once again, biochemistry and enzymology led to breakthroughs in chemical analysis.

Chemical Degradation and End-Labeling Oligonucleotide sequencing began with a typical chemical analysis for very short products of chemical oligonucleotide synthesis, for instance, chromatographic separations with different solvents before and after partial chemical degradation in a reaction specific for cleavage of chemical bonds in a base-specific manner (e.g., purine/pyrimidine specific). The amounts of chemicals involved were relatively large (mgs). Flat acrylamide gels were used for separating the reaction products. As noted above, the first RNA sequence, gene sequence, and total viral genome sequence were determined *without* the aid of gene cloning. However, it was not until DNA cloning and R.endo cleavage allowed the isolation of specific DNA fragments that sequencing could be used widely for gene analysis [178, 179]. End-labeling with P^{32} phosphate was dependent on the fact that oligonucleotide dephosphorylase and oligonucleotide kinases had been identified, produced, and purified for wide distribution (see Table 1.5). It was characteristic of this period, before companies had assembled a supply of materials for the burgeoning rDNA laboratories, that each group would produce a few key enzymes or vectors themselves and exchange them for other materials, such as host strains and vectors with other laboratories, both at home and abroad. International restrictions on rDNA, as well as commercial interests in it and its products, later reduced this practice severely. These halcyon days transformed to "Golden Helix" and "Genome Wars" to name just two book titles that describe the 1980–1990s entry of molecular genetics into the commercial world.

Oligonucleotide Primer Extension on a Single-Stranded Template (Sanger Sequencing) Initially, (radionuclide) end-labeled primers were extended in four separate base-specific terminating synthesis reactions in the presence of small amounts of strand-terminating dideoxynucleotide triphosphates, in addition to the normal four deoxynucleotide triphosphates (dT, dA, dG, and dC; dNTPs), a method developed by Sanger and Alan Coulson in 1977 [160, 194, 195]. These two scientists were also driving forces in developing the sequencing center at Cambridge, which made a major contribution to the human genome-sequencing

Table 1.5 Genetic engineering and sequencing required pure well-characterized enzymes.

Enzyme	Application	Discovered/references
Alkaline phosphatase (Calf thymus)	Removal of terminal 5′-phosphates before $P^{32}O_4$-labeling of oligonucleotides. Prevention of insert DNA scrambling	Morten [250]
Polynucleotide kinase (PNK)	End-labeling oligos (synthetic or dephosphorylated) (many sequencing methods; tracking hybridization etc.)	Novogrodsky *et al.* [180]
Haelll restriction endonuclease (representative of hundreds)	Preparing specific DNA fragments with blunt ends Specific cleavage of SV40 DNA by Hae R. endos.	Smith and Wilcox [181] Danna and Nathans [182]
EcoRI restriction endonuclease (representative of hundreds)	Preparing specific DNA fragments with cohesive ends	Boyer, [252] Mertz and Davis [183]
Terminal transferase (TdT)	Adding oligonucleotide tails to ds oligonucleotides (joining DNA fragments; cloning cDNA). Detecting apoptotic cells	Bollum [251]
DNA ligase from (*E. coli*/T4 ϕ)	Covalent joining DNA fragments with paired cohesive ends	Gellert [255] and Goulian and Kornberg [254]
Exonuclease I	Trimming single-stranded tails from ds DNA	Lehman and Nussbaum [184]
Exonuclease III (λphage)	Preparing 3′ ssDNA overhangs from 5′ PO_4 dsDNA (pre-tailing)	Sriprakash *et al.* [185]
DNA polymerase I (*Klenow fragment*)	Efficient DNA synthesis on a template (Lacking 3′ – 5′ exonuclease)	Lehman *et al.* [253] Klenow and Henningsen [186]
Reverse transcriptase (AMV)	Allows synthesis of a DNA strand complementary to an mRNA	Temin and Mizutani [187] Baltimore [188]
RNAse H	Digests RNA in RNA/DNA hybrids	Stein and Hausen [189]
S1 single-strand nuclease	Digests single-stranded DNA and ssDNA extension on dsDNA Analysis of DNA/DNA or DNA/RNA heteroduplexes	Ando [190] Vogt [191]
Bal31 ss endonuclease	Progressive deletion from ends of dsDNA (making deletions)	Gray *et al.* [192]
T. aquaticus DNA polymerase Heat stable	PCR (polymerase chain reaction) DNA amplification – basis for second-generation sequencing	Chien *et al.* [193] Lawyer *et al.* [154]

program (now called *The Sanger Center*). Gel columns were used later and formed the basis of separation in Leroy Hood's 1986 automated sequencing prototype that was developed as a workhorse for the final stages of the Human Genome Project in 1999.

Sequencing Automation: Large-Scale Sequencing Methods differ in several parameters such as the length of a sequence (length) that is reliably produced without error (raw error rate), the overall speed (raw base sequence per second), and cost. The increase of speed and reduction in cost of sequencing was, for each, over a million-fold in the last 25 years, a rate of development that is more dramatic than that proposed by "Moore's law" for computer technology: Moore's Law predicts doubling in speed and cutting cost by half every 2 years. The rate constant for sequencing is closer to doubling and halving, respectively, every year. As previously discussed, the development of PCR (particularly emulsion PCR as in the 454 sequencers using pyrosequencing), purification of relevant enzymes used in sequencing reactions, novel chemistries, such as reversibly terminating deoxynucleotides with different colored fluorescence for each type of base, and laser detection methods such as microarray scanning, were all important prerequisites for one or other of the sequencing technologies. The mega-sequencing technology, for example, used in the Illumina and Applied Biosystems, about 2008, simultaneously scans millions of sequencing reactions of randomly spaced "clonal amplicons" immobilized on a surface. Although each sequence length is short, the rate of sequencing more than compensates.

Next-Generation Sequencing Pettersson *et al*. [196] provide an excellent review into the previous and ongoing development of brilliant novel principles and their application in the several generations of sequencing technology up to 2009. Sequencing on single molecules can deliver long sequences as the method is devoid of the rising background noise generated by accumulated errors from each step of stochastic sequencing procedures involving many copies of each sequence.

In January 2016, one DNA sequencer supplier provides a machine that ostensibly delivers 1800 Gbp of sequence from 6000 million reads, each read giving 2×150 bp per run. This is far more than is necessary for sequencing an entire human genome to a very high degree of accuracy. The next paradigm shift required is a new method to deal with the data flood. New scientific journals are being created to deal with this new dimension, for example, "GigaScience" was founded 2011 as an open access journal.

1.5.6 DNA: Its Transfer to and Selection in Living Cells

In 1943, Avery, MacLeod, and McCarty had demonstrated that the transfer of DNA (and not RNA, protein, or lipids) was correlated with inheritance of an altered property (phenotype) in the recipient Streptococcal cells (pneumococci), an effect which could be seen both at the level of colony morphology (polysaccharide production when the cells were grown on solid surfaces) and

in terms of pathogenicity of the cells when injected into mice. This stimulated microbiologists and biochemists to investigate DNA in more detail culminating in X-ray crystallography. DNA crystal X-ray diffraction data produced by Rosalind Franklin were prematurely used by Watson and Crick. They built a tentative model of the DNA structure, communicated in a one-page letter to Nature in April 1953. The model cartooned a molecule of symmetrical beauty implying that it could be suitable as a carrier for genetic (inherited) information in the form of a long linear code. The presence of an antiparallel strand was interpreted as being a suitable template that would allow replication, mutation, and repair. The model structure was very close to the truth in its essential features and was rapidly accepted. In the 1970s, DNA transformation in which double-stranded supercoiled DNA was "transformed" into bacteria was established and led to the observation that these plasmids were often carrying antibiotic resistance genes [80, 81].

A parallel development of equal importance to Avery and co-workers in 1943 was the attempt to repeat this "transformation" by DNA transfer to animal cells. To demonstrate this, Szybalski made nonreverting mutants in the purine synthetic pathway (hypoxanthine-guanine phosphoribosyltransferase gene; HGPRT a salvage enzyme preventing xanthine and uric acid synthesis) of human cell lines, so that expression of the active *hgprt*-gene allowed selection, positive or negative, on a special medium (HAT medium). These mutants were, in fact, similar to those discovered later for the Lesch–Nyhan syndrome. Transferring DNA from normal cells to the mutant cell lines allowed selection of stable "wild-type" cells, the DNA from which, in turn, could "transform" other mutant cell lines [197]. This was a breakthrough that opened up human and other eukaryotic genetic analysis *in vitro*, and allowed selection of cell fusions. For example, such *hybridomas* were used to first produce mABs or cell lines carrying one, or a limited number of, human chromosomes. One could in fact perhaps in retrospect reevaluate the importance of this finding, saying that genetic engineering really began at this point in 1963!

The production of selectable vectors for use in animals cells often used Szybalski's system or further developments based on selective protocols directed to the purine synthesis pathway, for example, Bacchetti and Graham [198], who transferred the *Herpes simplex* thymidine kinase into a human cell line, albeit with very low efficiency in 1977; and Mulligan and Berg [199] who developed a selection protocol for the expression of an *E. coli gpt* gene in a eukaryotic vector in 1981.

1.5.7
Gene Cloning (1971–1982) the Era of Modern Biotechnology Based on Molecular Biology Begins

1.5.7.1
A Prerequisite for Cloning: Nucleic Acid, Biochemistry, and Enzymology

The possibilities to label, manipulate, and modify DNA molecules relied to a great extent on the availability of the pure, well-characterized enzymes listed in Table 1.5.

1.5.7.2

Applying Known Methods or a Conceptual Jump: the Details

Plasmids and viruses, particularly, have their own complex multicomponent structures (pilli and viral capsids) for transferring their DNA into host cells, with very limited host range. Often, the plasmids and viruses themselves carried DNA modification and destruction systems to inhibit competition and/or kill their hosts. Lateral DNA transfer (conjugation, transduction) between related bacteria was known, often promiscuous and widespread if strong selection pressure such as antibiotic treatment was applied. In the absence of such selection, it was rarely detectable. The spread of antibiotic resistance was, however, a main cause of concern among hospital staffs and a main topic of research for microbiologists including Stanley Falkow and Stanley Cohen in California who were among the first to have the idea of gene cloning (see Section 1.4). The discussion between Boyer and Cohen in the presence of three others, including Falkow, is assumed to be the crystallization point in the realization of gene cloning and the direct precursor to Boyer, who in 1976, with the venture capitalist Swanson, founded Genentech. At least for Cohen and Falkow, it seemed obvious that the isolation of specific DNA fragments (restriction fragments) and their overproduction in pure form (preparation of the recombinant plasmids from the clones) was a simple way to analyze the structure and function of the plasmids, which were causing promiscuous spread of antibiotic resistance in pathogens [80, 81, 148, 200–203]. The jump to cross species barriers with the aim of making pharmaceutical products was probably more in the forefront of Boyer's thoughts, an approach which allowed him to convince investors, in particular the risk capitalist Swanson, that protein therapeutics could be produced in bacteria. This was *contrary to the general conception of the difficulties that would have to be overcome* as outlined in the following list of the state of the art at the time.

Imponderables to Cross-Species Gene Transfer and Expression Among the imponderables that implied how impracticable it would probably be to move and stably express genes across species barriers, were the following:

- It was not all that clear that the genetic code was completely universal (it is not).
- Protein secretion is often mediated by protein–membrane complexes that recognize specific target proteins.
- There are hundreds of types of post-translational modification (proteolytic processing, addition of chemical groups, side chains) of proteins, often determining protein immunogenicity, function, solubility, and so on. This area was largely unexplored.
- Small peptides are scavenged and degraded in bacteria. *Note*: the statement that somatostatin was the first peptide (14aa) gene to be cloned and expressed in *E. coli* is not technically correct. A large chimeric protein had already been expressed, isolated from the bacteria, and chemically cleaved to yield the peptide fragment [204].
- Little was known about protein folding. Chaperone proteins are often necessary for correct folding of proteins and subunit assembly. In bacteria, there is low

oxygen concentration where disulfide bridges do not form. This is in contrast to the oxidizing condition in the cytoplasm of eukaryotes.
- The highly ordered folding of DNA (chromosomes) only existed in eukaryotes. Its role in gene expression was unknown.
- In eukaryotes, regulated secretion, sequestration to and protein release from organelles, as well as mRNA stabilization, were not understood.
- Some DNA structures are unstable, being able to jump to other locations (deletion and insertion) or invert (discovered in plants in 1951 by Barbara McClintock); rediscovered in bacteria, insertion elements, transposons, homing nucleases, CRISPR, and so on, targeted and destabilized, particularly foreign, DNA.
- Sequences with inverted mirror symmetry, that is, palindromes are extremely unstable (unclonable) in prokaryotes. The stability of direct repeats also differs.
- GC methylation common in eukaryotic DNA leads to DNA degradation in most *E. coli* hosts.
- Genetic load was not understood. Recombinant organisms are normally at a severe disadvantage in a natural environment. Under strong selection, recombinants can be forced to survive in the laboratory. However, often the rDNA is still rapidly lost, for example, where high transcription rates are involved.

With all the above in mind, the general impression in 1971 was that gene transfer, controlled gene expression, and stable inheritance, particularly across species barriers, would not be likely or would lead to transient-crippled unstable hybrids with poor expression that would certainly not be suitable as a source for biotechnology products. The mainstream of the classical pharmaceutical industry management was of this opinion, right up to the announcement that Genentech's recombinant insulin had been approved for clinical use in 1984.

By the end of 1974, it was understood that gene cloning was a generally applicable methodology, at least for academic laboratories *to investigate gene structure and function*. In 1977, DNA sequencing started to spread as a general tool. Armies of researchers battled to discover and deal with the practical consequences of the barriers presented above, many turning to species-specific cloning and expression systems, others continuing to deal with overcoming interspecies expression and protein-folding problems. The complexity of post-translational protein modification such as glycosylation has really only been solved by using tissue culture with cells from higher organisms for production for important pharmaceuticals, such as human tPA, EPO, the IFNs, and mABs [72].

How Gene Cloning Started? It has become commonplace to cite the conceptual beginning of gene cloning as being the 1972 paper from Paul Berg's laboratory [205]. As Berg himself admits [148], the paper in fact contains no novel methodology and most of the key enzymes were given to the group as gifts. It should be pointed out that, at that time, whether or not a method would work depended very much on the purity of the enzymes used, none of which was commercially available. The paper *neither demonstrates the feasibility of using the proposed*

protocol for transferring DNA to a foreign organism, nor is this idea original. The key principle in the Berg-Lobban method that was declared as novel was to join two oligonucleotides via overlapping hybridization, filling out single-strand region, and ligating with DNA ligase. This had been, in fact, demonstrated in Khorana's laboratory as a step in the synthesis of a completely synthetic gene for a tRNA [176]. The oligonucleotide elongation with Sgaramella's TdT (terminal deoxyribonucleotidyl transferase) as proposed in Berg's paper does not, in fact, work very well in practice for cloning purposes and leads to loss of the original restriction sites, thus complicating further analysis, fragment isolation, sequencing, and/or recloning (see discussion on p. 127 and Figure 1.1 in [155]). It did not, therefore, find much application initially (although it was later attractive as a step in complementary DNA (cDNA) cloning), the simpler and efficient method of Cohen and Boyer [206] being the obvious choice. A key discovery in Berg's laboratory was the closed circular ligation of EcoRI cleaved DNA by Metz that implied that the cohesive ends of the DNA were all the same and probably small and cohesive, melting in the range 4–15 °C, a fact which Boyer's laboratory (which had donated the EcoRI) immediately confirmed by DNA sequencing [207]. Rich Roberts particularly followed the route of Hamilton Smith and Boyer with respect to searching for and characterizing further restriction enzymes and using them in analyzing viral and plasmid DNA structure leading, for instance, to his discovery of gene splicing [208].

Danna and Nathans [182] were the first to cleave SV40 with one of the first type II restriction enzymes. Berg's group made *in vitro* SV40 hybrids with phage DNA; however, they transferred it *neither* to *E. coli nor* to eukaryotic cells, leaving open the question as to whether or not opening the restriction site destroyed function(s) required for the vector replication or gene expression. It was left finally to others to use restriction enzymes to investigate SV40 gene structure and function (e.g., the transformation gene, T, t) after transformation of cleaved viral fragments into an animal cell line [209]). A widely used early protocol for transfer of DNA into animal cell cells was coprecipitation of the DNA with calcium phosphate [210].

In 1976, Goff and Berg [211] selected for an *E. coli gpt* and a yeast *tRNA* gene cloned in an SV40 vector that could replicate in monkey cells. These clones could be developed as vectors for further studies of genes in eukaryotic cells using the foreign genes as selective markers. This was the first demonstration of cloning of bacterial DNA with ensuing cross-species transfer into animal cells. Initially, two types of SV40 vectors were available, (i) those that contained only the SV40 replication origin and could carry some 4–5 kb of foreign DNA and required cocultivation with helper virus and (ii) those with small deletions in the "late gene region" that could carry only small foreign fragments.

Transformation of cells by uptake of foreign DNA into bacteria was known since 1944. It was unknown whether eukaryotic DNA could be expressed in microorganisms, but the assumption that the chance of this happening might not be zero caused Berg and his group to postpone attempting cross-species transformation with their *in vitro*-ligated DNA. Perhaps this direction of thought was kept alive by ghosts of irreproducible work published in 1964 by, for example,

Thomas Trautner on Polyoma virus replication and production in transformed *Bacillus subtilis* (see review in [212]). The discussion initiated by Berg led to an excited meeting at Asilomar and the moratorium on gene cloning until guidelines for safe design of such experiments were established [148]. A more detailed analysis of the state of the art might, however, have arrived at a different conclusion that, initially, the most that could be expected from cloning eukaryotic DNA would be improved deletion analysis (specific induced deletion) of viruses and plasmids without expecting cross-species gene expression, the latter being the feeding ground for horror scenarios and public anxiety.

1.5.8 Genome Mapping: Clone Libraries, Restriction Maps, and RFLPs

1.5.8.1 Prelude: Human Genetics before Genome Sequencing

Barbara McClintock was incredibly far ahead of her time: she was one of the main developers of cytogenetics and the originator of the idea of chromosome meiosis and epigenetic effects in plants in the 1940s. She admonished students in life sciences, first of all, to "know your organism." Having a sequence does not imply that one has any knowledge relevant to your organism. Standing alone, it gives little or no insight and initially one has no idea where genes might be. It is only through the pioneering work of hundreds of human geneticists such as Victor McKusick and Leena Peltonen-Palotie, as well as biochemists and cell biologists who worked with them, that sequence work on the human genome could later be interpreted meaningfully in terms of the causality of heritable diseases.

The work of several generations of biochemists working with physicians insured that by 1973, 93 inborn errors of metabolism had been listed. Most of these detailed discoveries of the biochemical pathways and the enzymes involved in metabolic disease were made between 1957 and 1973 [213].

1.5.8.2 Important Ideas on Gene Mapping in the DNA Age

As DNA clones became available, it was considered how these might be used as "probes" or "links" to locate heritable disease genes on the chromosome. Again, it was McClintock, with her studies of meiosis in 1931, who had shown how the relative position of genes could be ordered along the chromosome. In 1980, a scheme was outlined as to how more-detailed linkage maps might be obtained for the whole human genome, using size variation of restriction fragments from the same region but from various individuals (restriction fragment length polymorphism*s*, RFLP*s*) [151]. This was also a central theme for the initial work of Jean Weissenbach at Genethon in Evry near Paris, founded in 1990 with a main agenda to develop diagnostics, and possibly treatments, for rare inherited diseases. One of the major breakthroughs that largely changed the negative attitude toward rDNA work was the identification of the dominant genes for the Lesch – Nyhan syndrome (where affected children die a terrible death at an early age). This knowledge was

applied as prenatal diagnostic to dramatically reduce the birth rate of affected children among the Ashkenazi Jews of New York. In 1993, the causal gene for Huntington Chorea, another incurable disease involving neuronal degeneration, was isolated and located to a chromosomal segment designated 4p16.3. This was the first autosomal disease locus found using genetic linkage analysis [214] and is associated with restriction fragment linked polymorphisms (RFLPs) where long stretches of triplet repeats in the gene are unstable. Since the symptoms appear late in life, long after child-bearing age, and are untreatable, (family) diagnostics came in for considerable criticism with the "right not to know" being in the forefront of the arguments.

A variety of developments in cloning technology improved the ability to isolate and analyze long regions of the chromosome in particular, simplifying the ordering of adjacent or overlapping clone fragments and assembly of clusters of overlapping sequences ("contigs"). Figures below in brackets give the number of primary clones needed to give a 95% confidence of covering the whole human genome:

- Initially, cloning was inefficient with plasmids rarely carrying fragments of greater than 5 kb with a bias more in the 1–2 kb range (10^7).
  Using the *lambda in vitro* packaging system, cloning became much more efficient and cloning of exclusively larger DNA fragments was ensured (see also [72]).
- with *lambda* cloning vectors, in which a "buffer" fragment had been deleted, inserts were in the size range of 5 kb, later 10–15 kb (6×10^5) [215]
- with *cosmids*; plasmids carrying the *cos*-site of lambda, which are packageable *in vitro* only if they have picked up at least 25–40 kb of insert DNA (2×10^5). Some designate low copy variants, with an F plasmid origin, as fosmids. The original vectors were amplifiable up to high copy number, facilitating further analysis and reducing contamination by *E. coli* chromosomal DNA [216].
- *M13 filamentous phage*. Viral cloning vectors can be isolated as a dsDNA plasmid-form from the cell and as ssDNA from the phage particles. It was developed for cloning by Messing *et al.* [217]. This latter can serve as template for sequencing by the Sanger primer elongation method [160, 218] using a synthetic DNA primer adjacent to the cloning site. It can also be applied to the introduction of specific mutations into the cloned gene by using a partially mismatched primer. It is the vector type used for developing phage-display technology, where clones can be physically selected on the basis of affinity to an immobilized target molecule of choice [219].

A powerful extension of this latter method was "sexual PCR" (DNA shuffling) developed by Stemmer [220] to (i) generate mutants and recombinants in mutant libraries, on a large scale and (ii) to select for novel ligands, which in combination with phage display, could be applied for ligand "affinity maturation" in analogy to the maturation of B-cells producing antibodies with improved affinity to specific target molecules. This recognized the importance of recombination in addition to point mutants in evolution, a factor also used most effectively for altering enzyme

specificity. (*Note*: this *in vitro* recombination system can also be used effectively to generate combinatorial library fragments that can also be selected *in vitro*, in combination with *in vitro* gene expression systems (e.g., [221].)

- Following some 9 years after Struhl's *et al.* [222] 1979 demonstration of cloning in yeast, Burke *et al.* [223] showed that very large fragments could be stably cloned in yeast (yeast auxiliary chromosome*s*; YAC*s*) with a size range of some 40–700 kb (average 80 kb); (10^5). Although reisolating the intact large insert from the yeast cell presented difficulties, YACs could be used for hybridization mapping and preparing sublibraries for sequencing and contig mapping. The libraries are very difficult to produce and are of limited use in tiling (joining or bridging contigs) because to up to 50% of the inserts are of chimeric origin. Their use was dropped in the latter part of the human genome project.
- *PAC vector DNAs* are packaged in a low-copy-number *E. coli* vector using an *in vitro* P1 phage packaging mix [224].

Each generation of cloning vector inventors praised the improved stability of their vector system. Much of this may be hearsay, although it may also be correlated with improvements made to the host strains, which allowed stabilization of the insert DNA and improved protocols, thus reducing palindrome formation of vectors while preparing the chimeric DNA. Purity of enzymes used in DNA manipulation also helped (e.g., purification of cleaner enzymes after cloning and expressing the gene for the enzyme; see Table 1.5).

- *Human accessory chromosome* (HAC) *vectors* , introduced in 1997 by Harrington *et al*. [225] demonstrated clones carrying up to 10 000 kb of DNA. These types of stable vectors are of more importance as gene transfer vectors to human cells to help annotate the human genome without disruption of the extant chromosomes and without use of potentially dangerous fragments of viral vectors.
- *Shuttle vectors* can be replicated and selected for in two hosts. *E. coli* is usually the host for initial isolation and manipulation of DNA as efficient cloning systems had been established for this host. It is then transferred to a new host by transformation or as in the case of *plants*, first into an *Agrobacterium* host, which then uses the T-DNA transfer system to transfer the DNA fragment to the plant host, for example, 1983 [226]. This latter system, the origin of rDNA work with plants, derived directly from studies of the growth of crown gall tumors where a natural system of cell transformation was discovered in which a tumorigenic DNA fragment was transferred from *A. tumefaciens*. Such vectors are of great use in expression cloning in prokaryotes (gene mining), particularly with respect to polyketide antibiotics; and primary and secondary metabolites.

Cosmids, BACs/PACs were used as the major resources in all the large-scale sequencing projects. Cosmids, particularly, are still used (sometimes as shuttle vectors) in "gene mining" (not the *in silico* form of gene mining which refers to searching for homology in databases) in the search for novel enzyme genes where expression of large genes or operons is required: they carry exclusively appropriate

sized DNA inserts, are easy to work with, and allow rapid mapping and sequencing of the cloned genes.

The presence of long repeated sequences throughout the genome makes it not only difficult for shotgun sequencing, it also makes the detection of overlapping fragments extremely difficult. Because of this, it was important to gather information from both ends of long sequences of defined length. With such data from, for example, "jumping and linking libraries," the relative position of two contigs separated by long repetitive sequences could be established. *Even today, there is still some 1% of the human genome sequence that is not established, although the regions in question are mapped to a few positions on the human chromosomes.*

- *Jumping and linking libraries*: For those involved for some 15 years in the1980s and 1990s in assembling a total gene/sequence map of large genomes, before total genome shotgun sequencing, a main task was to place a particular fragment to be sequenced with respect to the rest of the known sequences. This yielded the groups of known overlapping sequences and clones called *contigs*. Bridging contigs, that is, finding out the relative position of the contigs separated by as yet nonsequenced regions, was accomplished by sequencing clone libraries, especially made to contain only the ends of very large size-fractionated fragments that had been circularized *in vitro* ("jumping and linking-libraries") [227–229].

Some libraries were made in plasmid vectors exclusively from cDNAs (synthesized on an mRNA template). Many of these contained the open-reading frame from a spliced mRNA. The sequences of these clones were referred to as expressed sequences (ESs) and were used late in the genome sequence program to preferentially locate possible gene region tags (EST; e.g., [230]).

Since the advent of total shotgun genome sequencing and application of superior computer algorithms and hardware for ordering single sequence stretches within very large data sets [231], the earlier methods of gene fragment ordering before the advent of DNA sequencing, have become largely obsolete. Particularly for organisms where "the" genome has already been sequencing, "resquencing" becomes simplified. The shotgun approach was used first to obtain the total genome sequence of a microorganism (*Haemophilus influenzae Rd* in 1995 [232]). The first complete shotgun sequence of a larger eukaryote genome was that of the *Drosophila melanogaster* genome published by Gerald Rubin and Craig Venter's group in 2000 [233].

The human genome, being more than 3000-fold larger than that of *H. influenza* and 25 times larger than that of *Drosophila*, is considerably more difficult to master, relying exclusively on shotgun sequencing. As Venter himself admits, although his human genome sequencing project at Celera was largely based on the shotgun approach, much faster for accumulating raw data, he had used data from contig-linkage mapping and the EST-linkage method which he himself had helped pioneer. In spite of this, his choice of a single genome (his own) for sequencing simplified the linkage in the repetitive regions that in his case would be far less heterogeneous than that provided by the general sequencing (HUGO)

community. The single-genome approach also gave for the first time the actual sequence for *each of the diploid chromosome copies separately*. This approach, that took place first in an industrial format with nonpublic funding at Celera under the leadership of Venter, was much maligned by notable leaders of the human genome program in America who had little empathy with Venter's withholding his data for several weeks or months before releasing it. Many in this competitive field found Venter's reveling in his ability to sequence faster and more cheaply than the rest of the community fairly unbearable. A deplorable aspect of this phase of the genome-sequencing project was the extreme infighting for funding and committee lobbying. Much of this attitude appeared to be a continuation of the extreme competitiveness already extant from the early "gene wars," that is, "the characterization of contestants in the cloning races – Goeddel and Gilbert, and many others in supporting roles, … portray a sorry picture of personal rivalries in the conduct of science. A scientist can be intensely competitive and even unscrupulous in pursuit of a laudable goal" ([83], pp. 196, 198, 199). Also recommended is [234] for a coverage of the early development of the scientific and social issues in the human genome project, covering the heroes and the hypocrites.

The final outcome of the Celera sequencing effort was cost effective. Critics say it scavenged public research data financed by the taxpayer (true or false; who benefitted most from whom?). Some argue that scientists should have been doing more scientific research rather than having *generations of students* doing rote sequencing. Watson's supporters still maintain that Venter could not have completed the sequence without the data from the open HGP database, and may be correct. Conversely, the sequence was also certainly "finished" (still bits missing?) faster (and cheaper) with the help of the independent approach at Celera. We note that as data-flow increases exponentially, emphasis shifts toward improved and more rapid analysis. A little-recognized milestone was the development of the Genome Assembler algorithm by Granger Gideon Sutton at Celera (a nonprofit Organization led by Craig Venter) which succeeded in 1995 in yielding a complete *H. influenzae* genome (an NIH grant proposal was refused), later the *D. melanogaster* genome, and finally the human genome in 1999 all in the face of aggressive skepticism by the Jim Watson-lead "Genome Consortium" (NCBI) that such a random ("shotgun") assembly was impossible (see also [235]).

We note once more that without the accompanying medical, biochemical, and cell biological research on the *biological function of a sequence*, the sequence alone has no intrinsic value. When Francis Collins was asked at a press conference on the completion of the DNA sequence of a particular chromosome about the special significance of this event he said "a chromosome is an entity"; Craig Venter pithily remarked " a piece of shit is also an entity." Why this flagrant breach of political correctness? Perhaps it was necessary to explain to the Press reporting this incident that the worldwide genome project relied on grouping sequence efforts to groups supplied with clones selected as belonging to a particular chromosome so as to later simplify ordering of the sequences generated, for example, to "contigs" of overlapping clones. Membership in such a group required accepting this regime and was a necessary part of the funding program of which Venter was not a part.

Francis Collins should have admitted that he no choice other than this approach. Venter on the other hand did.

1.5.8.3

DNA Hybridization Chip Sequencing and More

Patrick O'Brown and Ronald Davis played a key role in the development of synthesis of oligonucleotide microarrays and their use for quantifying hybridization with nonradioactive oligos (optical sensors) that became a laboratory tool starting in about 1995. Once a genome has been sequenced in its entirety, it can be scanned for polymorphisms by direct hybridization to a DNA array (chips) carrying nucleotide primers corresponding to known polymorphic sites (sites with known sequence divergence) within PCR-amplified regions. The nucleotides hybridizing to the probe oligos on the chip are fluorescent labeled. This methodology has a very wide range of applications, for example, detecting GCNVs; routine screening for particular viruses or antibiotic resistance genes; whole-genome resequencing; comparative DNA methylation studies; distribution of epigenetically modified chromatin on the DNA (combined with chromatin immunoprecipitation, ChIP); and comparative gene transcription studies among others (section 11.4.5 in [72]). A combination of these latter lead to a better understanding of how genes respond to their environment, as well as ashow cells differentiate during embryogenesis, and how this can be reversed when cancer cells are placed with embryonic tissue.

1.5.8.4

Mega Sequencing: Impact on Biotechnology

The early phase of human genome sequencing gave rapid progress in tracking down the genes involved in many of the thousands of heritable diseases, syndromes, or traits that had been identified by human geneticists. Many of the powerful *biological response modifiers*, such as IFN, lymphokines, and growth and cell-differentiation factors, were first clearly characterized as proteins via sequencing and expressing their genes. This provided, at the same time, methods to reproducibly produce these compounds relatively cheaply in large amounts. This changed medical research and medical diagnostics, and increased the range of biopharmaceutical products (see Section 1.4).

The initial impact related to *first-generation products* such as insulin and growth hormone, also blood-clotting factors, that were already well characterized and already in clinical use. Their production via rDNA methods in bacteria or yeast gave an alternative source for a product that no longer had to be derived from human or animal blood or tissue and wase thus free of the danger of contaminating viruses from such sources (e.g., cytomegalovirus, hepatitis viruses, and AIDS).

First Pathogen-Free Source of Pharmaceuticals A fact that is sometimes overlooked in reviews of this period is the identification, cloning, and sequencing of pathogenic viruses. This was not just of general interest. It allowed for the first time the development of DNA-based PCR testing of products derived from blood

or tissues for the presence or absence of these viruses ([72], pp. 141–142; eight examples between 1979 and 1990, e.g., [236]) including the AIDS HIV virus. It also lead to the first rDNA-based vaccine for hepatitis B.

The second generation of products could not be characterized in detail until their genes had been cloned and they had been produced in pure form, and in sufficient amounts to be made available for detailed research, and eventually, in several cases, clinical application. One example is that of GCSF (granulocyte colony stimulating factor), cloned at Amgen in 1986 [237], which could be used since the late 1980s for blood stem-cell transplantation to cure, for instance, chemotherapy-resistant Hodgkin-lymphoma. Without this compound to stimulate blood stem-cell release into the donor blood, this treatment would not be an option (the youngest son of one of the authors of this article is still alive, thanks to this development; see Section 1.4 for more examples).

Welte [238] provided detailed insight into the academic, industrial, and medical collaborations that were involved and required at that time to complete this success story, starting with recognition of a protein function, purification of the protein, partial sequence, use of this information to find a cDNA clone sequence synthesized on a messenger RNA (mRNA), cloning the entire gene, and expression and optimization of production and purification, finally leading to extensive clinical testing (see Table 1.6 for methods involved).

The human genome program and the major advances in high-throughput sequencing has influenced biotechnology in a number of ways:

Table 1.6 Methods for enriching mRNAs/detecting specific clones or gene products.

System	Application	Developed by
Southern blotting (DNA) and Northern blotting (RNA)	Nucleic acid transfer to cellulose: detect homology to DNA or RNA probes (colony or phage plaque hybridization similar: → clone identification)	Southern [256] "Southern" Alwine *et al.* [257] "Northern"
Western blotting (1D or 2D)	Transfer of protein to filter: identify with monoclonal antibody	Symington *et al.* [258]
Hybridoma technology	First method of producing monoclonal antibodies (hybridomas not stable)	Köhler and Milstein [259]
rDNA monoclonal antibodies	Produced in cell lines, yeast, insect larvae, or bacteria (more stable production)	For example, Riechmann *et al.* [260]; Marks *et al.* [261]
Fluorophor- and enzyme-labeling	Protein or nucleic acid tagging with various highly fluorescent tags (e.g., GFP) many applications: DNA sequencing, tissue ID, chromosome painting (mapping) replaces radioactivity in the methods above. Also, enzyme-linked signal amplification (ELISA)	For example, review: Phillips [262]; Lichter, P *et al.* [263]; Ju *et al.* [264]

Diagnosis for Predisposition to Disease The *diagnostic possibilities for predisposition* to common diseases are of more limited clinical relevance than many originally expected. This has limited the predicted immediate growth of personal medicine. Sequence-based clinical diagnostics for the some 4000 hereditary mostly recessive traits known before the HUGO sequencing effort had really begun can now be carried out rapidly and with high precision without involving other family members. However, medical family history relating to inherited disorders (anamnesis) will always play a major role in indicating if a genetic causality should be investigated. Preventive surgery in the case of actress Angelina Jolie, who had a very high predisposition to breast cancer associated with particular BRCA gene alleles (over 70% likelihood rather than the general population 5–12%), has highlighted this diagnostic option in which sequence patenting has also played a historic role.

Genomewide Association Sequencing: Genes Affecting Widespread Chronic Syndromes When referring to "the human genome sequence," one does not usually refer to an individual sequence (apart from those of Venter and later of Watson). "The Sequence" was, and still is, the sum of data from thousands of genomes and hundreds of thousands of partial sequences of variations in particular disease-associated regions. The database provides *the basis for future medical and fundamental research related to human health*. However, many investigators warn against putting too strong a bias on looking for genetic factors relating to cancer, for instance, where the impact has been disappointing. By 2010, the US National Cancer Institute's TCGA had cost $375 million; part of an international cancer genome consortium that aims to sequence 25 000 tumor samples for a total of $1 billion. As stated by Robert Weinberg (MIT), "Sequencing endless more cancer genomes isn't going to tell us more than we already know" (http://news.sciencemag.org/2010/04/updated-skeptic-questions-cancer-genome-projects).

1.5.9 Expressing Genes in Other Organisms: Transgenic Animals Carrying rDNA

This review has glossed over many areas that deserve more detailed description with respect to biotechnological development. One of these is the area of recombinant antibodies. This is an exciting area that has produced some of the most innovative products to be derived from rDNA technology: particularly in bacteria; yeasts, both intracellular and surface expression; in Baculovirus vectors in insect cells; in animal cell lines, and in transgenic animals. These studies involved both affinity maturation and realization of the end-product for clinical use. The successful development of this area required finding solutions to nearly all of the problems encountered in the development of genetic engineering as presented above. These aspects have all been discussed in some detail with a listing of all the clinical applications up to and including 2010, including the history, the science, and the business [72].

The production of transgenic animals carrying rDNA in their genome arose from a combination of advances in embryology [239], understanding the differentiation of stem cells [240] in the developing embryo, and development of viral vectors [241]. This has had a broad range of applications, from production of "knockout" mice for fundamental research to animal models for human disease (e.g., a mouse model for HCV infections [242]).

1.5.10 Future Trends

In the search for new therapeutic ligands, the healthy competition between protein modeling and empirical selection procedures continues (e.g., [243]). This is true also in the search for therapeutic targets where "systems biology" should contribute a more global view of the interplay of intracellular pathways using *in silico* techniques [244]. Synthetic biology provides a new tool, independent of existing structures, where novel enzymes or gene regulators are constructed *de novo* by assembly of "Lego"-like functional building blocks [245]. This latter complements the ongoing refinement of genetic engineering with respect to direct gene replacement or mutation in various organisms. In the area of industrial biotechnology, the science continues to develop as a sustainable alternative to chemical synthesis.

The discovery of the key factors (often referred to as the *Yamanaka factors*) which induce differentiated cells to form pluripotent stem cells ((PiPS – protein-induced pluripotent stem cells) e.g., from adult human fibroblasts) opens up the vision of creating collections of immortalized stem-cell cultures of a huge range of tissue specificity as potential source for medical application over and above the requirement for helping postirradiation or chemotherapy recovery of leukemia patients, or even to use this method as a form of personalized medicine [246]. Although this development was achieved through the application of rDNA technology, the final PiPS technology, a radical advance for biotechnology, no longer requires it.

The latest developments in the CRISPR/Cas9 methods of *in vivo* genetic engineering, represent a great increase in the ability to accurately manipulate genetic material by DNA replacement (recombination) directly in living cells or whole organisms, efficiently and with little or no unwanted genetic side effects. This is expected to revolutionize much of modern medicine and biotechnology in the near future (*Note:* the first genetically engineered animal, salmon, engineered to grow faster and intended for food, has just been approved by FDA) [247a, b]. One innovative example with CRISPR/Cas9 technology can be seen in the work of, for example, George Church's group in which this new method was used to destroy 69 retroviral copies in the genome of a pig in a single experiment. This is the first step to generate pigs which are free of retroviruses that might be harmful to man. Further altering tissue specific antigens could lead to swine capable of serving as a safe harbor for human organs. In North America, some 125 000 patients per year are in need of organ donors [248]. In passing, one

can refer to George Church (MIT) as the epitomy of the modern molecular geneticist, who has fully embraced the role of the entrepreneur, being founder of at least 9 genome-related companies and is associated with or advising some 81 companies.

From the foregoing, we see that new tools are developed opening novel opportunities for investigation, often replacing older methods. Investigators, who are patient and those who are not, have made major contributions. Isolated individuals and those who build huge consortiums, those believing in entrepreneurship and those who do not, have all played their roles. Recognizing the way to progress is a privilege of the prepared mind independent of political or peer ranking. Let us keep reporting with honesty.

In the quest for solutions, we can be assured to stumble upon surprises as biological systems have a memory and a potential complexity that exceeds the possibility of the physical matter of the Universe to give each of them substance.

References

1. McGovern, P.E. *et al.* (2004) Fermented beverages of pre- and proto-historic China. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 17593–17598.
2. Bud, R. (1993) *The Uses of Life: A History of Biotechnology*, Cambridge University Press.
3. Lovitt, R.W., Kim, B.H., Shen, G.-J., and Zeikus, J.G. (1988) Solvent production by microorganisms. *Crit. Rev. Biotechnol.*, **7**, 107–186.
4. Mermelstein, L.D., Papoutsakis, E.T., Petersen, D.J., and Bennett, G.N. (1993) Metabolic engineering of *Clostridium acetobutylicum* ATCC 824 for increased solvent production by enhancement of acetone formation enzyme activities using a synthetic acetone operon. *Biotechnol. Bioeng.*, **42**, 1053–1060.
5. Qureshi, N., Li, X.-L., Hughes, S., Saha, B.C. *et al.* (2006) Butanol production from corn fiber xylan using *Clostridium acetobutylicum*. *Biotechnol. Progr.*, **22**, 673–680.
6. Bremus, C., Herrmann, U., Bringer-Meyer, S., and Sahm, H. (2006) The use of microorganisms in L-ascorbic acid production. *J. Biotechnol.*, **124**, 196–205.
7. Xu, A., Yao, J., Yu, L., Lv, S. *et al.* (2004) Mutation of *Gluconobacter oxydans* and *Bacillus megaterium* in a two-step process of L-ascorbic acid manufacture by ion beam. *J. Appl. Microbiol.*, **96**, 1317–1323.
8. Saito, Y., Ishii, Y., Hayashi, H., Imao, Y. *et al.* (1997) Cloning of genes coding for L-sorbose and L-sorbosone dehydrogenases from *Gluconobacter oxydans* and microbial production of 2-keto-gulonate, a precursor of L-ascorbic acid, in a recombinant *G. oxydans* strain. *Appl. Environ. Microbiol.*, **63**, 454–460.
9. Perlman, D. (1979) in *Microbial Technology*, 2nd edn, vol. 1 (eds J.H. Peppler and D. Perlman), Academic Press, New York, pp. 521–527.
10. Koizumi, S., Yonetani, Y., Maruyama, A., and Teshiba, S. (2000) Production of riboflavin by metabolically engineered *Corynebacterium ammoniagenes*. *Appl. Microbiol. Biotechnol.*, **51**, 674–679.
11. Perkins, J.B., Sloma, A., Hermann, T., Theriault, K. *et al.* (1999) Genetic engineering of *Bacillus subtilis* for the commercial production of riboflavin. *J. Ind. Microbiol. Biotechnol.*, **22**, 8–18.
12. Huembelin, M., Griesser, V., Keller, T., Schurter, W. *et al.* (1999) GTP cyclohydrolase II and 3.4-dihydroxy-2-butanone 4-phosphate synthase are the

rate-limiting enzymes in riboflavin synthesis of an industrial *Bacillus subtilis* strain used for riboflavin production. *J. Ind. Microbiol. Biotechnol.*, **22**, 1–7.

13. Scott, A.I. and Roessner, C.A. (2002) Biosynthesis of cobalamin (vitamin B12). *Biochem. Soc. Trans.*, **30**, 613–620.
14. Demain, A.L., Daniels, H.J., Schnable, L., and White, F. (1968) Specificity of the stimulatory effect of betaine on the vitamin B12 fermentation. *Nature*, **220**, 1324–1325.
15. Benninga, H. (1990) *A History of Lactic Acid Making*, Kluwer Academic Publishers.
16. Bai, D.-M., Zhao, X.-M., Li, X.-D., and Xu, S.-M. (2004) Strain improvement and metabolic flux analysis in the wild-type and a mutant *Lactobacillus lactis* strain for L(+)-lactic acid production. *Biotechnol. Bioeng.*, **88**, 681–689.
17. Zhou, S., Causey, T.B., Hasona, A., Shanmugam, K.T. *et al.* (2003) Production of optically pure D-lactic acid in mineral salts medium by metabolically engineered *Escherichia coli* W3110. *Appl. Environ. Microbiol.*, **69**, 399–407.
18. Roehr, M. (1998) A century of citric acid fermentation and research. *Food Technol. Biotechnol.*, **36**, 163–171.
19. Hannan, M.A., Rabbi, F., Faizur Rahman, A.T.M., and Choudhury, N. (1973) Analyses of some mutants of *Aspergillus niger* for citric acid production. *J. Ferment. Technol.*, **51**, 606–608.
20. Miall, L.M. (1978) in *Economic Microbiology: Primary Metabolites*, vol. 1 (ed. A.H. Rose), Academic Press, New York, pp. 47–119.
21. Mattey, M. (1992) The production of organic acids. *Crit. Rev. Biotechnol.*, **12**, 87–132.
22. Fukaya, M., Tayama, K., Tamaki, T., Tagami, H. *et al.* (1989) Cloning of the membrane-bound aldehyde dehydrogenase gene of *Acetobacter polyoxogenes* and improvement of acetic acid production by use of the cloned gene. *Appl. Environ. Microbiol.*, **55**, 171–176.
23. Wang, D.I.C., Fleischaker, R.J., and Wang, G.Y. (1978) A novel route to the production of acetic acid by fermentation. *AIChE Symp. Ser.*, **74**, 105–110.
24. Parekh, S.R. and Cheryan, M. (1994) High concentrations of acetate with a mutant strain of *C. thermoaceticum*. *Biotechnol. Lett.*, **16**, 139–142.
25. Cheryan, M., Parekh, S., Shah, M., and Witjitra, K. (1997) Production of acetic acid by *Clostridium thermoaceticum*. *Adv. Appl. Microbiol.*, **43**, 1–33.
26. Sonoyamana, T., Tani, H., Matsuda, K., Kageyama, B. *et al.* (1982) Production of 2-keto-L-gulonic acid from D-glucose by two-stage fermentation. *Appl. Environ. Microbiol.*, **43**, 1064–1069.
27. Lynd, L.R. (1990) Large scale fuel ethanol from lignocellulose: potential, economics, and research priorities. *Appl. Biochem. Biotechnol.*, **24/25**, 695–719.
28. Ingram, L.O., Conway, T., Clark, D.P., Sewell, G.W., and Preston, J.F. (1987) Genetic engineering of ethanol production in *Escherichia coli*. *Appl. Environ. Microbiol.*, **53**, 2420–2425.
29. Doran, J.B. and Ingram, L.O. (1993) Fermentation of crystalline cellulose to ethanol by *Klebsiella oxytoca* containing chromosomally integrated *Zymomonas mobilis* genes. *Biotechnol. Progr.*, **9**, 533–538.
30. Lynd, L.R. (1989) Production of ethanol from lignocellulosic materials using thermophilic bacteria: critical evaluation of potential and review. *Adv. Biochem. Eng.*, **38**, 1–52.
31. Demain, A.L., Newcomb, M., and Wu, J.H.D. (2005) Cellulase, clostridia and ethanol. *Microbiol. Mol. Biol. Rev.*, **69**, 124–154.
32. Eggeling, L. and Sahm, H. (1999) in *Metabolic Engineering* (eds S.Y. Lee and E.T. Papoutsakis), Marcel Dekker, New York, pp. 153–176.
33. Kitano, K., Sugiyama, Y., and Kanzaki, T. (1972) L-Glutamate fermentation with acetic acid by an oleic acid requiring mutant. II. Inhibitory factors against the extracellular accumulation of L-glutamate. *J. Ferment. Technol.*, **50**, 182–191.

34. Nakao, Y., Kikuchi, M., Suzuki, M., and Doi, M. (1972) Microbial production of L-glutamic acid by glycerol auxotrophs and production of L-glutamic acid from *n*-paraffins. *Agric. Biol. Chem.*, **36**, 490–496.
35. Laneelle, G. and Clement, Y. (1986) Glutamate excretion mechanism in *Corynebacterium glutamicum*. 5th International Symposium on the Genetics of Industrial *Microorganisms*, Part B, pp. 247–252.
36. Eggeling, L. and Sahm, H. (2001) The cell wall barrier of *Corynebacterium glutamicum* and amino acid efflux. *J. Biosci. Bioeng.*, **92**, 201–213.
37. Nampoothiri, K.M., Hoischen, C., Bathe, B., Moeckel, B. *et al.* (2002) Expression of genes of lipid synthesis and altered lipid composition modulates L-glutamate efflux of *Corynebacterium glutamicum*. *Appl. Microbiol. Biotechnol.*, **58**, 89–96.
38. Demain, A.L., Jackson, M., Vitali, R.A., Hendlin, D. *et al.* (1965) Production of xanthosine-5′-monophosphate and inosine-5′-monophosphate by auxotrophic mutants of a coryneform bacterium. *Appl. Microbiol.*, **13**, 757–761.
39. Demain, A.L., Jackson, M., Vitali, R.A., Hendlin, D. *et al.* (1966) Production of guanosine-5′-monophosphate and inosine-5′-monophosphate by fermentation. *Appl. Microbiol.*, **14**, 821–825.
40. Demain, A.L. (1968) Production of purine nucleotides by fermentation. *Prog. Ind. Microbiol.*, **8**, 35–72.
41. Tseng, Y.H., Ting, W.Y., Chou, H.C., Yang, B.Y. *et al.* (1992) Increase of xanthan production by cloning *xps* genes into wild-type *Xanthomonas campestris*. *Lett. Appl. Microbiol.*, **14**, 43–46.
42. Dawes, E.A. (1988) Polyhydroxybutyrate: an intriguing biopolymer. *Biosci. Rep.*, **8**, 537–547.
43. Pedros-Alio, C., Mas, J., and Guerrero, R. (1985) The influence of poly-β-hydoxybutyrate accumulation on cell volume and buoyant density in *Alcaligenes eutrophus*. *Arch. Microbiol.*, **143**, 178–184.
44. Datta, R., Tsai, S.P., Bonsignore, P., Moon, S.H. *et al.* (1995) The technological and economic potential of poly(lactic acid) and its derivatives. *FEMS Microbiol. Rev.*, **16**, 221–231.
45. Foster, W. and Raoult, A. (1974) Early description of antibiosis. *J. R. Coll. Gen. Pract.*, **24** (19), 889–894.
46. Bentley, R. (2001) Bartolomeo Gosio, 1863–1944: an appreciation. *Adv. Appl. Microbiol.*, **48**, 229–250.
47. Gratia, J.P. (2000) André Gratia: a forerunner in microbial and viral genetics. *Genetics*, **156**, 471–476.
48. Raper, K.B. (1994) The development of improved penicillin-producing molds. *Ann. N.Y. Acad. Sci.*, **48**, 41–56.
49. Dixon, B. (2006) Sulfa's true significance. *Microbe*, **1**, 500–501.
50. Demain, A.L. and Masurekar, P.S. (1974) Lysine inhibition of in vivo homocitrate synthesis in *Penicillium chrysogenum*. *J. Gen. Microbiol.*, **82**, 143–151.
51. Friedrich, C.G. and Demain, A.L. (1977) Effects of lysine analogs on *Penicillium chrysogenum*. *Appl. Environ. Microbiol.*, **34**, 706–709.
52. Fawcett, P.A., Usher, J.J., and Abraham, E.P. (1976) in 2nd International Symposium on the Genetics of Industrial Microorganisms (ed. K.D. MacDonald), Academic Press, New York, pp. 129–138.
53. Kohsaka, M. and Demain, A.L. (1976) Conversion of penicillin N to cephalosporin(s) by cell-free extracts of *Cephalosporium acremonium*. *Biochem. Biophys. Res. Commun.*, **70**, 465–473.
54. Yoshida, M., Konomi, T., Kohsaka, M., Baldwin, J.E. *et al.* (1978) Cell-free ring expansion of penicillin N to deacetoxycephalosporin C by *Cephalosporium acremonium* CW-19 and its mutants. *Proc. Natl. Acad. Sci. U.S.A.*, **75**, 6253–6257.
55. Banko, G., Demain, A.L., and Wolfe, S. (1987) α-(L-α-aminoadipyl)-L-cysteinyl-D-valine synthetase (ACV synthetase): a multifunctional enzyme with broad substrate specificity for the synthesis of penicillin and cephalosporin precursors. *J. Am. Chem. Soc.*, **109**, 2858–2860.

56. Hollander, I.J., Shen, Y.Q., Heim, J., Demain, A.L. *et al.* (1984) A pure enzyme catalyzing penicillin biosynthesis. *Science*, **224**, 610–612.
57. Nagarajan, R., Boeck, L.D., Gorman, M., Hamill, R.L. *et al.* (1971) β-Lactam antibiotics from *Streptomyces*. *J. Am. Chem. Soc.*, **93**, 2308–2310.
58. Stapley, E.O., Jackson, M., Hernandez, S., Zimmerman, S.B. *et al.* (1972) Cephamycins, a new family of β-lactam antibiotics. I. Production by actinomycetes, including *Streptomyces lactamdurans* sp. n. *Antimicrob. Agents Chemother.*, **2**, 122–131.
59. Bost, P.E. and Demain, A.L. (1977) Studies on the cell-free biosynthesis of β-lactam antibiotics. *Biochem. J.*, **162**, 681–687.
60. Brown, A.G. (1986) Clavulanic acid, a novel β-lactamase inhibitor-a case study in drug discovery and development. *Drug Des. Dev.*, **1**, 1–21.
61. Kahan, F.S., Kahan, F.M., Goegelman, R.T., Currie, S.A. *et al.* (1979) Thienamycin, a new β-lactam antibiotic. I. Discovery, taxonomy, isolation, and physical properties. *J. Antibiot.*, **32**, 1–12.
62. Berdy, J. (1995) *Proceedings of the 9th International Symposium on the Biology Actinomycetes, Part 1*, Allerton Press, New York, pp. 3–23.
63. Shier, W.T., Rinehart, K.L., and Gottlieb, D. (1969) Preparation of four new antibiotics from a mutant of *Streptomyces fradiae*. *Proc. Natl. Acad. Sci. U.S.A.*, **63**, 198–204.
64. McArthur, H.A.I. (1997) The novel avermectin, Doramectin – A successful application of mutasynthesis. Abstract S13 of the 5th International Conference on the Biotechnology of Microbiology (Products: Novel Pharmacology Agrobiology Activities) Williamsburg, VA, 1997, p. 20.
65. Okanishi, M., Suzuki, K., and Umezawa, H. (1974) Formation and reversion of streptomycete protoplasts: cultural condition and morphological study. *J. Gen. Microbiol.*, **80**, 389–400.
66. Hopwood, D.A. (1988) Towards an understanding of gene switching in *Streptomyces*, the basis of sporulation and antibiotic production. *Proc. R. Soc. London*, **235**, 121–138.
67. Hopwood, D.A., Malpartida, F., Kieser, H.M., Ikeda, H. *et al* (1985) Production of hybrid antibiotics by genetic engineering. *Nature*, **314**, 642–644.
68. Floss, H.G. (2006) Combinatorial biosynthesis – potential and problems. *J. Biotechnol.*, **124**, 242–257.
69. Dougherty, T.J. and Barrett, J.F. (2001) ABT-773: a new ketolide antibiotic. *Expert Opin. Invest. Drugs*, **10**, 343–351.
70. Nichterlein, T., Kretschmar, M., and Hof, H. (1996) RP 59500, a streptogramin derivative, is effective in murine listerosis. *J. Chemother.*, **8**, 107–112.
71. Petersen, P.J., Jacobus, N.V., Weiss, W.J., Sum, P.E. *et al.* (1999) In vitro and in vivo antibacterial activities of a novel glycylcycline, the 9-*t*-butylglycylamido derivative of minocycline (GAR-936). *Antimicrob. Agents Chemother.*, **43**, 738–744.
72. Buchholz, K. and Collins, J. (2010) *Concepts in Biotechnology: History, Science and Business*, Wiley-VCH Verlag GmbH, Weinheim.
73. Buchholz, K., Kasche, V., and Bornscheuer, U. (2012) *Biocatalysts and Enzyme Technology*, Wiley-VCH Verlag GmbH, Weinheim.
74. Bud, R. (1994) in *Resistance to New Technology* (ed. M. Bauer), Cambridge University Press, Cambridge, pp. 293–309.
75. Dechema (1974) *Biotechnologie*, Dechema, Frankfurt.
76. Buchholz, K. (1979) in *Geplante Forschung* (eds W. van den Daele, W. Krohn, and P. Weingart), Suhrkamp Verlag, Frankfurt, pp. 64–116.
77. Jasanoff, S. (1985) Technological innovation in a corporatist state: the case of biotechnology in the Federal Republic of Germany. *Res. Policy*, **14**, 23–38.
78. OTA (US Office of Technology Assessment) (1984) Impact of Applied Genetics, OTA, Washington, DC.
79. Hotchkiss, R.D. (1979) in *The Origins of Modern Biochemistry* (eds P.R. Srinivasan, J.S. Fruton, and J.T. Edsall),

New York Academy of Sciences, New York, pp. 321–342.
80. Cohen, S.N. (2013) DNA cloning: a personal view after 40 years. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 15521–15529.
81. Falkow, S. (2001) I'll have the chopped liver please, or how I learned to love the clone. A recollection of some of the events surrounding one of the pivotal experiments that opened the era of DNA cloning. *ASM News*, **67**, 555–559.
82. Hall, S.S. (1987) *Invisible Frontiers*, Tempus Books of Microsoft Press.
83. Kornberg, A. (1995) *The Golden Helix*, University Science Books, Sausalito, CA.
84. Lear, J. (1978) *Recombinant DNA: The Untold Story*, Crown Press, New York.
85. Ullrich, A. (1980) Genentech-story. *Nachr. Chem. Tech. Lab.*, **28**, 726–730.
86. Behrendt, U. (2009) *Genentech, Data, Technologietransfer-Workshop*, VBU, Dechema, Frankfurt.
87. Walsh, G. (2007) *Pharmaceutical Biotechnology*, John Wiley & Sons, Ltd., Chichester.
88. (a) Lähteenmäki, R. and Lawrence, S. (2005) Public biotechnology 2004 – the numbers. *Nat. Biotechnol.*, **23**, 663–671; (b) Lähteenmäki, R. and Lawrence, S. (2006) Public biotechnology 2005 – the numbers. *Nat. Biotechnol.*, **24**, 625–634.
89. (a) C&EN: Chemical & Engineering News (2009) vol. 7 (July 6), pp. 30–46; (b) July 4, (2011) pp. 34–39.
90. Roche (2008) *Geschäftsbericht 2008*, F. Hoffmann La-Roche AG, Basel.
91. Läsker, K. (2008) Sueddeutsche Zeitung, 19 May 2008, p. 20; 27 May 2008, p. 23.
92. Ernst & Young (2001) Yearly Report cited in the VCI (Verband der chemische Industrie), Yearly Report (Jahresbericht), Frankfurt, p. 24.
93. Chemical & Engineering News (2015) (Sept. 7), pp. 14–20.
94. Chemical & Engineering News (2015) C&E Supplement, pp. 26–29.
95. Chemical & Engineering News (2015) (June 29), pp. 20–21.
96. Heinzle, E., Biwer, A.P., and Cooney, C.L. (2006) Development of Sustainable Bioprocesses, Modelling and Assessment, John Wiley & Sons, Inc., New York.
97. Weuster-Botz, D., Hekmat, D., Puskeiler, R., and Franco-Lara, E. (2007) Enabling technologies: fermentation and downstream processing. *Adv. Biochem. Eng./Biotechnol.*, **105**, 205–247.
98. (a) Walsh, G. (2005) in *Modern Biopharmaceuticals*, vol. 1 (ed. J. Knaeblein), Wiley-VCH Verlag GmbH, Weinheim, pp. 1–34; (b) Wurm, F.M. (2004) Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat. Biotechnol.*, **22**, 1393–1398.
99. Zhou, T.-C., Zhou, W.-W., Hu, W., and Zhong, J.-J. (2010) *Encyclopedia of Industrial Biotechnology*, John Wiley & Sons, Inc., New York, http://tinyurl.com/wileyEIB (accessed 31 May 2016).
100. Deckwer, W.-D., Jahn, D., Hempel, D., and Zeng, A.-D. (2006) Systems biology approaches to bioprocess development. *Eng. Life Sci.*, **6**, 455–469.
101. Hempel, D.C. (2006) Development of biotechnological processes by integrating genetic and engineering metods. *Eng. Life Sci.*, **6**, 443–447.
102. Klein-Marcuschamer, D., Yadav, V.G., Ghaderi, A., and Stephanopoulos, G.N. (2010) De novo metabolic engineering and the promise of synthetic DNA. *Adv. Biochem. Eng./Biotechnol.*, **120**, 101–131.
103. Silverman, J. *et al.* (2005) Multivalent avimer proteins evolved by exon shuffling of a family of human receptor domains. *Nat. Biotechnol.*, **23**, 1556–1561.
104. Chemical & Engineering News (2015) (Nov. 9), p. 9.
105. Papini, M., Salazar, M., and Nielsen, J. (2010) Systems biology of industrial microorganisms. *Adv. Biochem. Eng./Biotechnol.*, **120**, 51–99.
106. Wittmann, C. (2010) Analysis and engineering of metabolic pathway fluxes in *Corynebacterium glutamicum* for amino acid production. *Adv. Biochem. Eng./Biotechnol.*, **120**, 21–49.
107. Marx, A., Eikmanns, B.J., Sahm, H., de Graaf, A.A., and Eggeling, L. (1999)

Response of the central metabolism in *Corynebacterium glutamicum* to the use of an NADH-98 Hiller, dependent glutamate dehydrogenase. *Metab. Eng.*, **1**, 35–48.

108. Pfefferle, W., Möckel, B., Bathe, B., and Marx, A. (2003) Biotechnological manufacture of lysine. *Adv. Biochem. Eng./Biotechnol.*, **79**, 59–112.

109. Raab, R.M., Tyo, K., and Stephanopoulos, G. (2005) Metabolic engineering. *Adv. Biochem. Eng./Biotechnol.*, **100**, 1–17.

110. Melzer, G., Eslahpazir, M., Franco-Lara, E., and Wittmann, C. (2009) Flux design: in silico design of cell factories based on correlation of pathway fluxes to desired properties. *BMC Syst. Biol.*, **3**, e120.

111. Soetaert, W. and Vandamme, E. (eds) (2010) *Industrial Biotechnology*, Wiley-VCH Verlag GmbH, Weinheim.

112. Elander, R.P. (2003) Industrial production of β-lactam antibiotics. *Appl. Microbiol. Biotechnol.*, **61**, 385–392.

113. RFA (Renewable Fuels Association) (2005) http://www.ethanolrfa.org/industry/locations/ (accessed 31 May 2016).

114. Bruggink, A. (2001) *Synthesis of ß-Lactam Antibiotics*, Kluwer Academic Publishers, Dordrecht.

115. Schulz, W.G. (2007) The costs of biofuels. *Chem. Eng. News*, **85** (51), 12–16 (emissions adapted form *Science* (2006) **311**, 506).

116. U.S. Department of Energy (DEO) (2004) *Energy Efficiency and Renewable Energy, Biomass: Top Value Added Chemicals from Biomass* (eds T. Werpy and G. Petersen), U.S. Department of Commerce, Springfield, VA.

117. Kamm, B. and Kamm, M. (2007) Biorefineries – multi product processes. *Adv. Biochem. Eng./Biotechnol.*, **105**, 175–204.

118. EU (2005) http://ec.europa.eu/research/energy/index_eu.htm.

119. Schröder, U., Niessen, J., and Scholz, F. (2003) A generation of microbial fuel cells with current outputs boosted by more than one order of magnitude. *Angew. Chem.*, **115**, 2986–2989.

120. Bornscheuer, U. and Buchholz, K. (2005) Highlights in biocatalysis – historical landmarks and current trends. *Eng. Life Sci.*, **5**, 309–323.

121. Tosa, T., Mori, T., Fuse, N., and Chibata, I. (1969) Studies on continuous enzyme reactions. 6. Enzymatic properties of DEAE-sepharose-aminoacylase complex. *Agric. Biol. Chem.*, **33**, 1047–1056.

122. (a) Mayer, H., Collins, J., and Wagner, F. (1979) Patent DE2930794; (b) McCain, K.W. (1995) The structure of biotechnology R & D. *Scientometrics*, **32** (2), 153–175.

123. Buchholz, K. and Poulson, P.B. (2000) in *Applied Biocatalysis* (eds A.J.J. Straathof and P. Adlercreutz), Harwood Academic Publishers, Amsterdam, pp. 1–15.

124. Andersen, O. and Poulsen, E.B. (1983) in *Enzyme Technology* (ed. R.M. Lafferty), Springer-Verlag, Berlin, pp. 179–188.

125. Klibanov, A. and Cambou, B. (1987) in *Methods in Enzymology*, vol. 136 (ed. K. Mosbach), Academic Press, Orlando, FL, pp. 117–137.

126. Cheetham, P.S.J. (2000) in *Applied Biocatalysis* (eds A.J.J. Straathof and P. Adlercreutz), Harwood Academic Publishers, Amsterdam, pp. 93–152.

127. Jördening, H.-J. and Winter, J. (2005) *Environmental Biotechnology*, Wiley-VCH Verlag GmbH, Weinheim.

128. Buchholz, K., Kasche, V., and Bornscheuer, U. (2005) *Biocatalysts and Enzyme Technology*, Wiley-VCH Verlag GmbH, Weinheim.

129. Hubschwerlen, C. (2007) in *Comprehensive Medicinal Chemistry II*, vol. 7 (eds J.J. Plattner and M.C. Desai), Elsevier, Amsterdam, pp. 497–517.

130. Jarvis, L.M. (2007) *Chem. Eng. News*, 15–20.

131. Aggarwal, S. (2007) What's fueling the biotech engine? *Nat. Biotechnol.*, **25**, 1097–1104.

132. Melmer, G. (2005) in *Production of Recombinant Proteins. Novel Microbial and Eukariotic Expression Systems* (ed. G. Gellissen), Wiley-VCH Verlag GmbH, Weinheim, pp. 361–383.

133. Aggarwal, S. (2014) What's fueling the biotech engine – 2012–2013. *Nat. Biotechnol.*, **32**, 32–39.
134. Jarvis, L.M. (2015) The year in new drugs. *Chem. Eng. News*, **93** (5), 11–16.
135. Chemical & Engineering News (2014) (Nov. 24), p. 6.
136. Slater, A., Scott, N.W., and Fowler, M.R. (2008) *Plant Biotechnology: The Genetic Manipulation of Plants*, Oxford University Press, Oxford.
137. Stephanopoulos, G. (1999) Metabolic fluxes and metabolic engineering. *Metab. Eng.*, **1**, 1–11.
138. Reuss, M. (2001) Editorial. *Bioprocess Biosyst. Eng.*, **24**, 1.
139. Buchholz, K. (2007) Science or not? The status and dynamics of biotechnology. *Biotechnol. J.*, **2**, 1154–1168.
140. Yi, D. (2011) Who owns what? Private ownership and the public interest in recombinant DNA technology in the 1970s. *Isis*, **102**, 446–474.
141. Zucker, L.G. and Darby, M.R. (1997) Present at the biotechnological revolution: transformation of technological identity. *Res. Policy*, **26**, 429–446.
142. Zucker, L.G., Darby, M.R., and Armstrong, J. (1998) Geographically localized knowledge: spillovers or markets? *Econ. Inq.*, **36**, 65–86.
143. Zucker, L.G. and Darby, M.R. (2001) Capturing technological opportunity via Japan's star scientists: evidence from Japanese firms' biotech patents and products. *J. Technol. Transfer*, **26**, 37–58.
144. Arts, S. and Veugelers, R. (2012) *The Technological Origins and Novelty of Breakthrough Inventions*, KU Leuven, http://ssrn.com/abstract=2230366 (accessed 31 May 2016).
145. Arts, S., Appio, F.P., and Van Looy, B. (2013) Inventions shaping technological trajectories: do existing patent indicators provide a comprehensive picture? *Scientometrics*, **97** (2), 397–419.
146. (a) Rheinberger, H.-J. (1997) *Toward a History of Epistemic Things*, Stanford University Press, Stanford, CA; Also:(b) Rheinberger, H.-J. (2001) *Experimentalsysteme und epistemische Dinge*, Wallstein–Verlag, Göttingen.
147. (a) Bartneck, C. and Rauterberg, M. (2007) Physics Nobels should favour inventions. *Nature*, **448**, 644; (b) Bartneck, C. and Rauterberg, M. (2008) The asymmetry between discoveries and inventions in the Nobel Prize in Physics. *Technoetic Arts*, **6**, 73.
148. Berg, P. and Mertz, J.E. (2010) Personal reflections on the origins and emergence of recombinant DNA technology. *Genetics*, **184**, 9–17.
149. Venter, J.C. (2008) *A Life Decoded: My Genome: My Life*, Penguin Press Science, New York.
150. Bassett, A.R. and Ji-Long Liu, J.-L. (2014) CRISPR/Cas9 and genome editing in Drosophila. *J. Genet. Genomics*, **41**, 7–19. doi: 10.1016/j.jgg.2013.12.004
151. Botstein, D., White, R.L., Skolnick, M. *et al.* (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.*, **32**, 314–331.
152. Saiki, R.K. *et al.* (1985) Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, **230**, 1350–1354.
153. Saiki, R.K. *et al.* (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, **239**, 487–491.
154. Lawyer, F.C., Stoffel, S., Saiki, R.K., Myambo, K., Drummond, R., and Gelfand, D.H. (1989) Isolation, characterization, and expression in *Escherichia coli* of the DNA polymerase gene from *Thermus aquaticus*. *J. Biol. Chem.*, **264**, 6427–6437.
155. Collins, J. (1977) Gene cloning with small plasmids. *Curr. Top. Microbiol. Immunol.*, **78**, 122–170.
156. Collins, J. (1997) in *Annual Reports in Combinatorial Chemistry and Molecular Diversity* (eds W.H. Moos, M.R. Pavia, B.M. Kay, and B.A. Ellington), Springer-Verlag, Berlin, New York, pp. 210–267, ISBN: 978-94-017-0738-1 (Print), 978-0-306-46904-6 (Online).
157. Winnacker, E.-L. (1987) *From Genes to Clones*, VCH Verlagsgesellschaft, Weinheim.
158. Berget, S.M., Moore, C., and Sharp, P.A. (1977) Spliced segments at

the 5' terminus of adenovirus 2 late mRNA. *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 3171–3175. PMCID: PMC431482.

159. Chow, L.T., Gelinas, R.E., Broker, T.R., and Roberts, R.J. (1977) An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell*, **12**, 1–8.
160. (a) Sanger, F. (1980) *Determination of Nucleic Acid Sequences in DNA*, Nobel Prize Lecture; (b) Shizuya, H., Birren, B., Kim, U.-J., Mancino, V., Slepak, T., Tachiira, Y., and Simon, M. (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-based vector. *Proc. Natl. Acad. Sci. U.S.A.*, **89**, 8794–8797.
161. Maher, B. (2008) Personal genomes: the case of missing hereditability. *Nature*, **456**, 18–21.
162. Gilbert, J.A., Jansson, J.K., and Knight, R. (2014) The earth microbiome project: successes and aspirations. *BMC Biol.*, **12** (1), 69.
163. de Duve, C. and Beaufay, H. (1981) A short history of cell fractionation. *J. Cell Biol.*, **91**, 293–299.
164. Pedersen, K.O. (1976) The development of Svedberg's ultracentrifuge, *Biophys. Chem.*, **5**, 3–18.
165. Shapiro, J., MacHattie, L., Eron, L., Ihler, G., Ippen, K., Beckwith, J., Arditti, R., Reznikoff, W., and MacGillivray, R. (1969) The isolation of pure *lac* operon DNA. *Nature*, **224**, 768–774.
166. Hirt, B. (1967) Selective extraction of polyoma DNA from infected mouse cell cultures. *J. Mol. Biol.*, **26**, 365–369.
167. Radloff, R., Bauer, W., and Vinograd, J. (1967) A dye-buoyant-density method for the detection and isolation of closed circular duplex DNA in HeLa cells. *Proc. Natl. Acad. Sci. U.S.A.*, **57**, 1514–1521.
168. de Chadarevian, S. (2002) Designs for Life. Molecular Biology after World War II, Cambridge University Press.
169. Brownlee, D.G. (2014) Fred Sanger Double Nobel Laureate: A Biography, Cambridge University Press.
170. (a) Holley, R.W. *et al.* (1965) Structure of a ribonucleic acid. *Science*, **147**, 1462–1465; (b) Holley, R.W. (1968) *Alanine Transfer RNA*, Nobel Prize Lecture.
171. Sanger, F. (1971) Nucleotide sequences in bacteriophage ribonucleic acid. The eighth hopkins memorial lecture. *Biochem. J.*, **124**, 833–843.
172. Brescher, M.S. (1968) Direct translation of a circular mRNA. *Nature*, **220**, 1088–1091.
173. Hindley, J., Staples, D.H., Billeter, M.A., and Weissmann, C. (1970) Location of the coat cistron on the RNA of phage Qß. *Proc. Natl. Acad. Sci. U.S.A.*, **67**, 1180–1187.
174. Min Jou, W., Haegeman, G., Ysebaert, M., and Fiers, W. (1972) Nucleotide sequence of the gene coding for the bacteriophage MS2 coat protein. *Nature (London)*, **237**, 82–86.
175. Fiers, W., Contreras, R., Duerinck, F., Haegeman, G., Iserentant, D., Merregaert, J., Min Jou, W., Molemans, F., Raeymaekers, A., Van den Berghe, A., Volckaert, G., and Ysebaert, M. (1976) Complete nucleotide-sequence of bacteriophage MS2-RNA – primary and secondary structure of replicase gene. *Nature*, **260**, 500–507.
176. (a) Kleppe, K., Ohtsuka, E., Kleppe, R., Molineux, I., and Khorana, H.G. (1971) Studies on polynucleotides *1, *2XCVI. Repair replication of short synthetic DNA's as catalyzed by DNA polymerases. *J. Mol. Biol.*, **56**, 341–361; (b) Khorana, H.G. (1979) Total synthesis of a gene. *Science*, **203**, 614–662.
177. Gilbert, W. and Maxam, A. (1973) The nucleotide sequence of the lac operator. *Proc. Natl. Acad. Sci. U.S.A.*, **70**, 3581–3584.
178. Maxam, A.M. and Gilbert, W. (1977) A new method for sequencing DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 560–564.
179. Gilbert, W. (1980) *DNA Sequencing and Gene Structure*, Nobel Prize Lecture.
180. Novogrodsky, A., Tal, M., Traub, A., and Hurwitz, J. (1966) The enzymatic phosphorylation of ribonucleic acid and deoxyribonucleic acid. II. Further properties of the 5'-hydroxyl polynucleotide kinase. *J. Biol. Chem.*, **241**, 2933–2943.
181. Smith, H.O. and Wilcox, K.W. (1970) A restriction enzyme from *Hemophilus*

influenzae. I. Purification and general properties. *J. Mol. Biol.*, **51**, 379–391.

182. Danna, K. and Nathans, D. (1971) Specific cleavage of simian virus 40 DNA by restriction endonuclease of *Hemophilus influenzae*. *Proc. Natl. Acad. Sci. U.S.A.*, **68**, 2913–2917.

183. Mertz, J.E. and Davis, R.W. (1972) Cleavage of DNA by RI restriction endonuclease generates cohesive ends. *Proc. Natl. Acad. Sci. U.S.A.*, **69**, 3370–3374.

184. Lehman, I.R. and Nussbaum, A.L. (1964) The deoxyribonucleases of *Escherichia Coli*. V. on the specificity of exonuclease I (Phosphodiesterase). *J. Biol. Chem.*, **238**, 2628–2636.

185. Sriprakash, K.S., Lundh, N., Huh, M.M.-O., and Radding, C.M. (1975) The specificity of lambda exonuclease. Interactions with single-stranded DNA. *J. Biol. Chem.*, **250**, 5438–5445.

186. Klenow, H. and Henningsen, I. (1970) Selective elimination of the exonuclease activity of the deoxyribonucleic acid polymerase from *Escherichia coli* B by limited proteolysis. *Proc. Natl. Acad. Sci. U.S.A.*, **65**, 168–175.

187. Temin, H.M. and Mizutani, S. (1970) RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature*, **226**, 1211–1213.

188. Baltimore, D. (1970) RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature*, **226**, 1209–1211.

189. Stein, H. and Hausen, P. (1969) Enzyme from calf thymus degrading the RNA moiety of DNA-RNA hybrids: effect on DNA dependent RNA polymerase. *Science*, **166**, 393–395.

190. Ando, T. (1966) A nuclease specific for heat-denatured DNA isolated from a product of *Aspergillus oryzae*. *Biochim. Biophys. Acta*, **114**, 158–168.

191. Vogt, V.M. (1973) Purification and further properties of single-strand-specific nuclease from *Aspergillus oryzae*. *Eur. J. Biochem.*, **33**, 192.

192. Gray, H.B. Jr., Ostrander, D.A., Hodnett, J.L., Legerski, R.J., and Robberson, D.L. (1975) Extracellular nucleases of *Pseudomonas BAL 31*. I. Characterization of single strand-specific deoxyriboendonuclease and double-strand deoxyriboexonuclease activities. *Nucleic Acids Res.*, **2**, 1459–1492.

193. Chien, A., Edgar, D.B., and Trela, J.M. (1976) Deoxyribonucleic acid polymerase from the extreme thermophile *Thermus aquaticus*. *J. Bacteriol.*, **127**, 1550–1557.

194. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 5463–5467.

195. García-Sancho, M. (2010) A new insight into Sanger's development of sequencing: from proteins to DNA, 1943–1977. *J. Hist. Biol.*, **43**, 265–323.

196. Pettersson, E., Lundeberg, J., and Ahmadian, A. (2008) Generations of sequencing technologies. *Genomics*, **93**, 105–111.

197. Szybalska, E.H. and Szybalski, W. (1962) Genetics of human cell lines. IV: DNA mediated heritable transformation of a biochemical trait. *Proc. Natl. Acad. Sci. U.S.A.*, **48**, 2026–2034.

198. Bacchetti, S. and Graham, F.L. (1977) Transfer of the gene for thymidine kinase to thymidine kinase-deficient human cells by purified herpes simplex viral DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 1590–1594.

199. Mulligan, R.C. and Berg, P. (1981) Selection for animal cells that express the *Escherichia coli* gene coding for xanthine-guanine phophoribyl transferase. *Proc. Natl. Acad. Sci. U.S.A.*, **78**, 2072–2076.

200. Silver, R.P. and Falkow, S. (1970) Specific labeling and physical characterization of R-factor deoxyribonucleic acid in *Escherichia coli*. *J. Bacteriol.*, **104**, 331–339.

201. (a) Hughes, S.S. (2009) Regional Oral History Office, Bancroft Library, University of California, Berkeley, CA; (b) Transcripts of Interviews with Cohen, S.N., Berg, P., Boyer, H., Falkow, S., and Heynecker, H. (1995) http://bancroft.berkeley.edu/ROHO/ (accessed 31 May 2016).

202. Cohen, S.N., Chang, A.C., and Hsu, L. (1972) Nonchromosomal antibiotic

resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **69**, 2110–2114.

203. Cohen, S.N. (1982) in *From Genetic Engineering to Biotechnology: The Critical Transition* (eds W.J. Whelan and S. Black), John Wiley & Sons, Inc., New York, pp. 213–216.

204. Itakura, K., Hirose, T., Crea, R., Riggs, A.D., Heyneker, H.L., Bolivar, F., and Boyer, H.W. (1977) Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin. *Science*, **198**, 1056–1063.

205. Jackson, D.A., Symons, R.H., and Berg, P. (1972) Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.*, **69**, 2904–2909.

206. Cohen, S., Chang, A., Boyer, H., and Helling, R. (1973) Construction of biologically functional bacterial plasmids *in vitro*. *Proc. Natl. Acad. Sci. U.S.A.*, **70**, 3240–3244.

207. Hedgpeth, J.H., Goodman, M., and Boyer, H.W. (1972) DNA nucleotide sequence restricted by the RI endonuclease. *Proc. Natl. Acad. Sci. U.S.A.*, **69**, 3448–3452.

208. Roberts, R.J. (2005) How restriction enzymes became the workhorses of molecular biology. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 5905–5908.

209. Abrahams, P.J., Mulder, C., Van De Voorde, A., Warnaar, S.O., and van der Eb, A.J. (1975) Transformation of primary rat kidney cells by fragments of simian virus 40 DNA. *J. Virol.*, **16**, 818–823.

210. Graham, F.L. and van der Eb, A.J. (1973) A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology*, **52**, 456–467.

211. Goff, S. and Berg, P. (1976) Construction of hybrid viruses containing SV40 and lambda phage DNA segments and their propagation in cultured monkey cells. *Cell*, **9**, 695–705.

212. Jones, D. and Sneath, P.H. (1970) Genetic transfer and bacterial taxonomy. *Bacteriol. Rev.*, **34**, 40–81.

213. Harris, H. (1975) *The Principles of Human Biochemical Genetics*, North-Holland Publishing Co., Amsterdam.

214. Macdonald, M. (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. *Cell*, **72**, 971–983, doi: 10.1016/0092-8674(93)90585-E. PMID 8458085.

215. Hohn, B. and Murray, K. (1977) Packaging recombinant DNA molecules into bacteriophage *lambda* particles *in vitro*. *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 3259–3262.

216. Collins, J. and Hohn, B. (1978) Cosmids: a type of plasmid gene-cloning vector that is packageable *in vitro* in bacteriophage λ heads. *Proc. Natl. Acad. Sci. U.S.A.*, **75**, 4242–4246.

217. Messing, J., Gronenberg, B., Müller-Hill, B., and Hofschneider, P.H. (1977) Filamentous coliphage *M13* as a cloning vehicle: insertion of a *HinII* fragment of the *lac* regulatory region in *M13* replicative form *in vitro*. *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 3542–3646.

218. Viera, J. and Messing, J. (1982) The pUC plasmids. An *M13mp7*-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene*, **19**, 259–268.

219. Smith, G.P. (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science*, **228**, 1315–1317.

220. Stemmer, W.P. (1994) DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc. Natl. Acad. Sci. U.S.A.*, **91**, 10747–10751.

221. Hannes, J. and Plückthun, A. (1997) *In vitro* selection and evolution of functional proteins by using ribosomal display. *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 4937–4942.

222. Struhl, K., Stinchcomb, D.T., Scherer, S., and Davis, R.W. (1979) High-frequency transformation of yeast: autonomous replication of hybrid DNA

molecules. *Proc. Natl. Acad. Sci. U.S.A.*, **76**, 1035–1039.

223. Burke, D.T., Carle, G.F., and Olson, M.V. (1987) Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science*, **236**, 806–812.

224. Ioannou, P.A., Amemiya, C.T., Garnes, J., Kroisel, P.M., Shiyuza, H., Chen, C., Batzer, M.A., and de Jong, P.J. (1994) A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nat. Genet.*, **6**, 84–89.

225. Harrington, J.J., Van Bokkelen, G., Mays, R.W., Gustashaw, K., and Huntington, W.F. (1997) Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. *Nat. Genet.*, **15**, 345–355.

226. Herrera-Estrella, L., Depicker, A., van Montagu, M., and Schell, J. (1983) Expression of chimaeric genes transferred into plant cells using a Ti-plasmid-derived vector. *Nature*, **303**, 209–213.

227. Collins, F.S. and Weissman, S.M. (1984) Directional cloning of DNA fragments at a large distance from an initial probe: a circularization method. *Proc. Natl. Acad. Sci. U.S.A.*, **21**, 6812–6816.

228. Poustka, A. and Lehrach, H. (1986) Jumping libraries and linking libraries: the next generation of molecular tools in mammalian genetics. *Trends Genet.*, **2**, 174–179.

229. Poustka, A., Pohl, T.M., Barlow, D.P., Frischauf, A.M., and Lehrach, H. (1987) Construction and use of human chromosome jumping libraries from NotI-digested DNA. *Nature*, **325**, 353–355.

230. Bailey, L.C. Jr., Searls, D.B., and Overton, G.C. (1998) Analysis of EST-driven gene annotation in human genomic sequence. *Genome Res.*, **8**, 362–376.

231. Staden, R. (1979) A strategy of DNA sequencing employing computer programs. *Nucleic Acids Res.*, **6**, 2601–2610.

232. Fleischmann, R.D., Adams, A.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A., Merrick, J.M. *et al.* (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae Rd*. *Science*, **269**, 496–512.

233. Adams, M.D. *et al.* (2000) The genome sequence of *Drosophila melanogaster*. *Science*, **287**, 2185–2195.

234. Kevles, D.J. and Hood, L. (eds.) (1992) *The Code of codes: Scientific and social issues of the human genome project*. Harvard University Press, Cambridge, Mass.

235. Levy, S. *et al.* (2007) The diploid genome of an individual human. *PLoS Biol.*, **5**(10), e254.

236. Wang, K.-S., Choo, Q.-L., Weiner, A.J. *et al.* (1986) Structure, sequence and expression of the hepatitis δ viral genome. *Nature*, **323**, 508–514.

237. Souza, L.M., Boone, T.C., Gabrilove, J. *et al* (1986) Recombinant human granulocyte colony stimulating factor: effects on normal and leukemic myeloid cells. *Science*, **232**, 61–65.

238. Welte, K. (2010) in 20 Years of GCSF: Clinical and Non-Clinical Studies Molineux (eds M. Foote and T. Arvedson), Springer-Verlag, Heidelberg, pp. 15–24.

239. Gordon, J.W. and Ruddle, F.H. (1981) Integration and stable germ line transformation of genes injected into mouse pronuclei. *Science*, **214**, 1244–1246.

240. Gossler, A., Doetschman, T., Korn, R., Serfling, E., and Kemler, R. (1986) Transgenesis by means of blastocyst-derived embryonic stem cell line. *Proc. Natl. Acad. Sci. U.S.A.*, **83**, 9065–9069.

241. Jaenisch, R. (1976) Germ-line integration and mendelian transmission of the exogenous Moloney leukemia virus. *Proc. Natl. Acad. Sci. U.S.A.*, **73**, 1260–1264.

242. Dorner, M., Horwitz, J.A., Robbins, J.B., Barry, W.T., Feng, Q., Mu, K., Jones, C.T. *et al.* (2011) A genetically humanized mouse model for hepatitis C virus infection. *Nature*, **474**, 208–211.

243. Kügler, J., Schmelz, S., Gentzsch, J., Haid, S., Pollmann, E., van den Heuvel, J., Franke, R., Pietschmann, T., Heinz, D.W., and Collins, J. (2012) High affinity peptide inhibitors of the hepatitis

C virus NS3-4A protease refractory to common resistant mutants. *J. Biol. Chem.*, **287**, 39224–39232.

244. Chuang, H.-Y., Matan Hofree, M., and Ideker, T. (2010) A decade of systems biology. *Annu. Rev. Cell Dev. Biol.*, **26**, 721–744.

245. Wang, Y.-H., Wei, K.Y., and Smolke, C.D. (2013) Synthetic biology: advancing the design of diverse genetic systems. *Annu. Rev. Chem. Biomol. Eng.*, **4**, 69–102.

246. Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, **131**, 861–872.

247. (a) Doudna, J.A. and Charpentier, E. (2014) The new frontier of genome engineering with CRISPR-Cas9. *Science*, **346**, 1077, doi: 10.1126/science.1258096; (b). Chemical & Engineering News (2015) (Nov. 23), p. 22.

248. Yang, B.L. *et al* (2015) Genome-wide inactivation of porcine endogenous retroviruses (PERVs). *Science*, **350**, 1101–1104.

249. Fulwyler, M.J. (1965) Electronic separation of biological cells by volume. *Science*, **150**, 910–911.

250. Morton, R.K. (1955) Some properties of alkaline phosphatase of cow's milk and calf intestinal mucosa. *Biochem. J.*, **60**, 573–582.

251. Bollum, F.J. (1960) Calf thymus polymerase. *J. Biol. Chem.*, **235**, 2399–2403.

252. Boyer, H.W. (1971) DNA restriction and modification mechanisms in bacteria. *Annu. Rev. Microbiol.*, **25**, 153–176.

253. Lehman, I.R., Bessman, M.J., Simms, E.S. and Kornberg , (1958) A. Enzymatic synthesis of deoxyribonucleic acid. I. Preparation of substrates and partial purification of an enzyme from Escherichia coli. *J. Biol. Chem.*, **233**, 163–170.

254. Goulian, M. and Kornberg, A. (1967) Enzymatic synthesis of DNA. 23. Synthesis of circular replicative form of phage *phi*-X174 DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **58**, 1723–1730.

255. Gellert, M. (1967) Formation of covalent circles of lambda DNA by *E. coli* extracts. *Proc Natl. Acad. Sci. U.S.A.*, **57**, 148–155.

256. Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.*, **98**, 503–517.

257. Alwine, J.C., Kemp, D.J. and Stark, G.R. (1977) Method for detection of specific RNAs in agarose gels by transfer to diazobenzoyloxymethyl-paper and hybridization with DNA probes. *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 5350–5354.

258. Symington, J., Green, M. and Brackmann, K. (1981) Immunoautoradiographic detection of proteins after electrophoretic transfer from gels to diazo-paper: analysis of adenovirus encoded proteins. *Proc. Natl. Acad. Sci. U.S.A.*, **78**, 177–181.

259. Köhler, G and Milstein, C. (1975) Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature (Lond.)*, **256**, 495–497.

260. Riechmann, L., Clark, M., Waldmann, H. and Winter, G. (1988) Reshaping antibodies for therapy. *Nature (Lond.)* **332**, 323–327.

261. Marks, J.D., Hoogenboom, H.R., Bonnert, T.P., McCafferty, J., Griffiths, A.D. and Winter, G. (1991) By-passing immunization: Human antibodies from V-gene libraries displayed on phage. *J. Mol. Biol.*, **222**, 581–597.

262. Phillips, G.J. (2001) Green fluorescent protein – a bright idea for the study of bacterial protein localization. *FEMS Microbiol. Lett.*, **204**, 9–18.

263. Lichter, P., Tang, C.J., Call, K., Hermanson, G., Evans, G.A., Housman, D. and Ward, D.C. (1990) High resolution mapping of human chromosome 11 by *in situ* hybridization with cosmid clones. *Science*, **247**, 64–69.

264. Ju J., Ruan, C., Fuller, C.W., Glazer, A.N. and Mathies (1995) Fluorescent energy transfer dye-labeled primers for DNA sequencing and analysis. *Proc. Natl. Acad. Sci. U.S.A.*, **92**, 4347–4351.