Selection markers and expression enhancing elements: novel strategies to create mammalian production cell lines
Hoeksema, F.

Citation for published version (APA):
Hoeksema, F. (2011). Selection markers and expression enhancing elements: novel strategies to create mammalian production cell lines

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
CHAPTER 1

GENERAL INTRODUCTION
1 BIOPHARMACEUTICALS

1.1 Introduction to biopharmaceuticals and historical background

Biopharmaceuticals are a collective name for therapeutic agents that are proteinaceous in nature and which are manufactured by biotechnology methods using living organisms or cells. Biopharmaceuticals structurally mimic compounds found within the body. In a more restricted definition, biopharmaceuticals are considered as therapeutic monoclonal antibodies (MAbs), therapeutic proteins, vaccines and nucleic acid based drugs. Since biopharmaceuticals are proteinaceous, they are mostly very large and they are commonly delivered by (intravenous) injection. This is all in contrast with so-called small molecules, therapeutic chemicals, which are produced by chemical synthesis. Since they are small, they are often suitable for oral administration. By and large, small molecule therapeutics have been at the basis of the successes of the pharmaceutical industry throughout the 20th century.

A few major developments can be discerned in the pharmaceutical industry that led to the big surge of biopharmaceuticals: 1) the advent of recombinant DNA technology in the 1970s, 2) the US Supreme Court decision that genetically modified organisms were patentable (Diamand v. Chakrabarty decision, 1980) and 3) the shift of attention of the pharmaceutical industry from small molecule therapeutics to biopharmaceuticals. The first biopharmaceuticals were purified from animal tissues. One of the best examples is insulin that was isolated from pig pancreases and which was used for the treatment of diabetes. Although successful, there were negative immunological side effects due to its animal origin. Biopharmaceuticals of human origin were also isolated. For instance, one of the world largest pharmaceutical companies Genzyme became successful with supplying a biopharmaceutical for the treatment of Gaucher disease, an otherwise fatal childhood disease. From human placentas, Genzyme isolated glucocerebrosidase, the enzyme that is defective in patients that have Gaucher disease. Minor quantities were isolated; 20,000 human placentas provided 1 year’s supply for a single patient. A final example is human growth hormone, which was isolated from the pituitary gland of human cadavers. Treatment with these hormones was connected to the development of Creutzfeldt-Jakob disease in about 100 patients worldwide. It was again insulin that in 1982 was approved by the US Federal Food and Drugs Administration (FDA) as the first human recombinant biopharmaceutical. It was and still is produced in bacteria by Ely Lilly, and with all types of variants, it is still good for an annual 1 billion US dollars, even after 28 years. In 1986, the FDA approved human tissue plasminogen activator (tPA), an antithrombotic agent, as the first biopharmaceutical that was produced from mammalian cells. Its manufacturer was Genentech, based in San Francisco. Genentech was one of the very first companies that was founded with the intention to use genes and biotechnology to produce therapeutic biopharmaceuticals. Genentech began to use large-scale batch cultures of up to 10,000 liter reactors of genetic engineered Chinese Hamster Ovary
(CHO) cells. Genentech has been a front-runner company ever since, being the first biotechnology company to create a marketing department for biopharmaceuticals. It practically invented biotechnology related patenting, both for specific products and for instance for methodologies to produce therapeutic proteins. Genentech was the first company to use mammalian cells as its production platform. Today, Genentech is still one of the most innovative biotech companies with most approved therapeutic biopharmaceuticals in the market.

The success of tPA paved the way for a significant number of other recombinant proteins. The next product to be licensed was erythropoietin (EPO), a hormone that controls the maturation of red blood cells. This product is also produced in genetically engineered CHO cells in a multi-stage roller-bottle process. The technique to synthesize EPO was patented by Columbia University and was licensed to Amgen in 1989. When Amgen launched Epogen it set a precedent that many of its competitors were to follow. That is, select a disease whose cause is a single human protein, has a large market and has no microbiologically produced or small molecule alternative.

Originally, during the first phase of the establishment of large pharmaceutical companies, a few so-called blockbuster small molecules were responsible for their initial tremendous growth. For instance, Aspirin was at the basis of success for Bayer, Librium and Valium for Hoffmann-La Roche (nowadays simply Roche), Zantac for GlaxoSmithKline. This initial phase of rapid growth during the beginning of the 20th century is long gone for a few reasons. In the first place, many blockbuster small molecule medicines have gone ‘off-patent’ for a long time. This has resulted in a still growing market for so-called generic medicines that are chemically identical to the often very expensive blockbuster medicines. Generic medicines are synthesized by different companies that sell these medicines at much lower prices. This is possible since the chemical synthesis of the small molecules has been described in detail in many regulatory documents. Also, these small molecules don’t have to be discovered again in the first place, and no elaborate testing in clinical trials is required prior to marketing, so the price can be much lower. Since the costs of health care have risen steeply over the past few decades, there is a lot of pressure on both physicians and insurance companies to prescribe generic medicines. In the second place it has turned out to be very difficult to find novel blockbuster medicines. Intensive screening efforts, along with ‘intelligent’ approaches such as proteomics to find new drug targets held the hope that a novel generation of small molecule medicines could be found. By and large these hopes have been false. This is one of the reasons that the ‘big pharma’ have turned their attention to biopharmaceuticals, even though this is not where their roots are.

Although highly profitable, biotechnology has remained a high risk enterprise. Huge investments have to be made in order to reach the point of putting a blockbuster therapeutic agent in the market. Even a successful company such as Genentech had to resort to an alliance with a ‘big pharma’ company, in this case Roche. Already in 1990, Roche took a 60% interest in Genentech. By this, in particular the sales of Genentech’s products had a more solid basis. By 2007, Genentech-developed therapeutic agents contributed to nearly a third of Roche’s sales. In particular, three MAbs against cancer, Avastin (anti-colon cancer), Rituximab (anti-non-Hodgkin lymphoma) and Herceptin (anti-breast cancer) were good for over 16 billion US dollars in 2009. The only recently (2004) approved Avastin alone sold 5,5 billion US dollars in 2009 and projections are that this will become 10 billion US dollars in 2015. Despite fierce opposition from Genentech itself, Roche bought the remaining 40% shares for over 45 billion US dollars in 2009. These figures demonstrate the immense value of biopharmaceuticals that has been created in just a few decades. And the Genentech-Roche story does not stand on its own. As of 2010, the ‘big pharma’ company Sanofi Aventis in France is involved in a (hostile) take-over of one of the largest biotechnology companies Genzyme, based in the USA. So far the 18.5 billion US dollars has been offered. Why is this trend so global?
Figure 1. Global biopharmaceutical market, with estimated 2009 revenues. Biopharmaceuticals can be divided into a number of major groups, classified by the nature/activity of the product and the commercial value of the group. The market values of those biopharmaceuticals are shown. ‘Proteins’ includes hematopoietic growth factors, blood factors, growth hormones and cytokines. Adapted from Evers, 2010.

The reasons for these take-overs are very similar as stated above. Major ‘big pharma’ industries suffer from loss of block bustar medicines. During the last two decades this has led to a consolidation in the pharmaceutical industry by fusions and take-overs. Currently less than 15 major ‘traditional’ pharmaceutical industries are left (Novartis, Pfizer, Bayer, GlaxoSmithKline, Johnson & Johnson, Sanofi-Aventis, Roche, AstraZeneca, Merck & Co., Abbott Laboratories, Wyeth, Eli Lilly and Company, Amgen and Boehringer Ingelheim). Buying often considerably smaller biotechnology companies ensures that the product pipelines of the pharmaceutical companies will be filled for the next years. After all, the current thread for biotechnology companies is that patents that protect their original products will be expired in the next years. As with small molecule medicines, this causes the emergence of so-called biosimilar therapeutic agents. MAbs and vaccines that are produced by mammalian cells can never be completely identical, in contrast to chemical synthesis of small molecules. Hence the term ‘biosimilar’ has been invented. Although not identical, the FDA demands only that the effects of the produced MAb or vaccine are the same in terms of efficiency and therapeutic result. This is vastly faster and cheaper to prove than to find a novel biopharmaceutical and hence a number of biotechnology-based companies have started to produce biosimilars. One aspect of this is that the so-called ‘cost of goods’ have to remain as low as possible. Hence, more efficient production methods are in need to achieve that.

Recombinant protein therapeutics have proven to be invaluable pharmaceuticals for the treatment of various diseases. The advent of recombinant biopharmaceuticals allowed for the production of proteins that are not immunogenic and risk for animal-origin or blood-borne disease is virtually eliminated. Much larger quantities of the biopharmaceuticals can be produced as opposed to the quantities that were originally isolated from animal or human sources. Despite these major improvements, increased production levels of biopharmaceuticals are still required. For instance, MAbs are administered at high doses, ranging from 25-40 mg every other week to 100 mg per day. Protein production at multi-kilogram scales to sustain such therapies is required. The advent of these therapies has triggered speculations about capacity shortages and has highlighted that even modest increases in product yield will significantly decrease the ‘cost of goods’.
High clinical dose requirements of those pharmaceuticals demand for large-scale, efficient and cost effective manufacturing processes. Thus, beside to reduce the ‘cost of goods’, also to meet these high demands, improvements in production are necessary. This is the major objective of the studies that are presented in this PhD thesis.

1.2 The biopharmaceutical market

Above, huge figures were quoted for the value of biopharmaceuticals. What are these figures based upon? Grossly, it can be said that MAbs dominate the market value, by representing roughly one-third of the global biopharmaceutical market. All therapeutic proteins together (including growth factors, blood factors, growth hormones and cytokines) represent half of the biopharmaceutical market and approximately 15% of the biopharmaceutical market is represented by vaccines (Figure 1). Analysts forecast growth in the biopharmaceuticals sector and a continued increase in the MAb’s share of the market, while vaccines hold their market share (i.e. increase as rapidly as MAbs) and therapeutic proteins, although growing steadily, are slight losers in market share (Table 1).

1.2.1 Monoclonal antibodies

When the structure of antibodies was discovered, it was soon realized that they could be used as therapeutic agents. The binding of antibodies to a target is highly specific, an aspect much desired in therapeutics. The advent of MAbs opened the road towards the creation of therapeutic MAbs. Traditionally, the most efficient way of producing antibodies was to immunize a large vertebrate, bleed the animal and, from the serum, collect the polyclonal protein. The extraction of tiny quantities of polyclonal antibodies and the immunogenicity of those animal-derived antibodies limited their applicability because a vast majority of the antibodies is non-neutralizing and every immunization will result in a new polyclonal mixture. Since the discovery of a method for producing antibodies from hybridomas, resulting from the fusion of B-lymphocytes with a myeloma, the production of relatively large amounts of MAbs was possible. The first MAbs were murine, but immunogenicity was still a problem and it had to be reduced. For example, a large fraction of the patients treated with OKT3, the first MAb approved by the FDA, produced human anti-mouse antibodies (HAMA). It turned out that in particular the mouse constant regions of the MAbs were responsible for the immunological reactions.

Hence, the solution became to remove these parts and replace them with human constant regions, creating chimeric MAbs. One accompanying positive effect was that the efficiency of these chimeric MAbs was also higher than of the corresponding mouse MAbs, which are

| Biopharmaceutical market, estimated value and forecast 2009-2015 ($bn) |
|-----------------------------|---|---|---|---|---|---|---|
| Year | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | CAGR, % |
| Mabs | 38 | 41 | 45 | 50 | 55 | 59 | 64 | 9.3 |
| Therapeutic proteins | 61 | 63 | 66 | 69 | 72 | 75 | 78 | 4.4 |
| Vaccines | 18 | 19 | 21 | 23 | 25 | 26 | 28 | 8.1 |
| Total market | 117 | 123 | 132 | 142 | 152 | 160 | 170 | 6.7 |

Table 1. Estimates and forecasts of the biopharmaceutical market for 2009-2015. Estimates and forecasts for the three main biopharmaceutical categories (MAbs, therapeutic proteins and vaccines) are given. The percentage CAGR (compound annual growth rate) is indicated. Adapted from Evers, 2010."
relatively ineffective due to their weak ability to recruit human immune system processes and their rapid removal from the human blood. Although immunological side effects are much less than with murine MAbs, chimeric MAbs can still exhibit unwanted immunological responses. Even more human are humanized MAbs, which are essentially identical to that of a human variant, except for some of its complementarity determining regions (CDR) segments of murine origin, which are responsible for the ability of the MAb to bind to its target antigen. Humanized antibodies exhibit minimal or no response of the human immune system against them. Finally, transgenic mice in which the mouse Ig loci are replaced with the human Ig loci and human antibody phage display libraries allow for the isolation of ‘fully human’ MAbs.

Today, most MAbs are humanized or ‘fully human’ versions and are produced in mammalian cells.

In terms of sales, MAbs covered 38% of the global biopharmaceutical market in 2009. As therapeutic agents, there are two major applications of MAbs, for the treatment of cancer and for the treatment of inflammatory and immune-related diseases. Antibodies can be used alone (naked antibodies) or joined to a chemotherapy drug, radioisotope or toxin (conjugated antibodies). Naked antibodies are the most commonly used and in cancer treatment where they often attach to malignant cells, they act as a marker so that the body’s immune system can recognize and destroy them. Some naked MAbs do not interact with a patient’s immune system, but attach to specific antigens that are functional parts of cancer cells. Whereas the idea of a ‘magic bullet’ was already postulated at the beginning of the 20th century, conjugated MAbs have been developed only recently and are used as delivery vehicles to take chemotherapy drugs, radioisotopes or toxins directly to cancer cells, lessening damage to healthy cells in other parts of the body. For the treatment of inflammatory and immune-related diseases, the common mode of action of the antibodies is that they target Tumor Necrosis Factor (TNF), a cytokine involved in systemic inflammation.

Within the MAb group, antibodies against various forms of cancer rank highest in sales. One of the first examples is trastuzumab (Herceptin), developed and marketed by Genentech, now Roche. Herceptin blocks the interaction between the epidermal growth factor (EGF) and the human EGF receptor 2 (HER2). This receptor is over expressed in certain types of invasive breast carcinomas and Herceptin has become a standard therapeutic agent to treat these subclasses of breast cancer. Although a huge ‘seller’ for Genentech with sales over 5 billion US dollars in 2009, the effects on the treatment of breast cancer have been not as hoped for. Only a limited population of breast cancer patients can be treated with Herceptin. In combination with chemotherapy, Herceptin prolongs survival of patients with HER2-positive breast cancer in metastatic and adjuvant settings. However, most patients with HER2-positive metastatic breast cancer eventually develop progressive disease. Consequently, additional therapeutic options for this patient population are needed. This has resulted in the development of next generation Herceptin-derived MAbs. For instance, T-DM1 is a novel anti-HER2 antibody-drug conjugate, which combines the HER2-targeting properties of Herceptin with intracellular delivery of maytansinoid (DM1), a cytotoxic agent. By the means of this conjugated MAb, exposure of HER2-positive tumors to DM1 is maximized, whereas exposure of normal tissue is minimized.

Beside highly successful MAbs against cancer, such as Herceptin and Avastin (that targets and inhibits the binding of vascular endothelial growth factor (VEGF) to its receptors) there are other examples of MAbs that have a large impact on human health care. Particularly MAbs against inflammation and immune-related diseases are high value therapeutic agents. For instance, adalimumab (Humira), produced by Abbott is a successful and commercially important MAb for the treatment of inflammatory diseases. It is a TNF inhibitor, by binding to TNFα, preventing it from activating TNF receptors. TNFα inactivation has proven to be important in down regulating the inflammatory reactions associated with
autoimmune diseases. Humira is a fully human MAb used for the treatment of rheumatoid arthritis, Crohn’s disease and other inflammatory disorders.

Whereas these MAbs target specific molecules and interfere with their function, other antibodies simply target an epitope on specific tumors, without interfering with the function of this epitope. Being attached to the tumor cell helps to evoke an immunological reaction from the patient, with as end result the destruction of the tumor cell by the immune system of the patient. An example of such antibodies is rituximab (Rituxan) (Genentech, now Roche), which targets CD20, an epitope marker that is present on more than 90% of B-cell lymphoma19. In fact, Rituxan was the first anti-cancer MAb that was commercially successful. It still had a ~15% share of the MAb market in 2009.

1.2.2 Non-MAb therapeutic proteins

Therapeutic proteins are proteins with a therapeutic application that directly bind a specific target via an interaction that is not an antigen-antibody bond. Therapeutic proteins were initially produced by chemical isolation from animal sources. For instance, growth hormone used to treat deficiencies was extracted from the pituitary glands of cadavers. Also, diabetes was originally treated with purified insulin derived from pig pancreases. Furthermore, specific blood proteins are of therapeutic use and those blood factors have traditionally been obtained from blood donated by human volunteers. However, since the recombinant DNA technology revolution of the 1970s, the arrival of ‘large molecule’ protein-based therapeutics was preluded. Recombinant DNA technology facilitated the industrial production of proteins whose absence or low levels in patients lead to disease. Present-day therapeutic proteins are exclusively generated using molecular biology. This allows for much larger quantities of therapeutic proteins available in contrast to the minor quantities that were isolated from animal sources. Moreover, immunogenicity reactions against the animal-derived proteins are avoided. Also, the risk of animal-origin or blood-borne disease is virtually eliminated. Finally, with recombinant DNA technology, therapeutic proteins with higher specific activity can be created.

Typically, therapeutic proteins are exogenously fabricated versions of extracellular signaling ligands and receptors. Examples of therapeutic proteins are insulin, blood factors, growth hormones, cytokines, and hemopoietic growth factors, in particular erythropoietin (EPO). In terms of sales, 61% of the global biopharmaceutical market was covered by therapeutic proteins in 2009, as compared to 38% by MAbs. The therapeutic proteins market in 2009 can be divided in sales by class: EPOs ($12.5 billion), Insulins ($12 billion), Coagulation factors ($6 billion), IFN-beta ($6 billion), G-CSF ($5.5 billion), IFN-alpha ($3 billion), hGH ($3 billion) and the remaining proteins ($13 billion).

EPO is a glycoprotein hormone that is normally produced by peritubular cells in the kidney and liver and it controls erythropoiesis, the formation of red blood cells. Industrially produced recombinant EPO20 is therefore used for the treatment of blood amnesia, for instance in patients with kidney problems and in particular in patients that have received chemotherapy21. EPO acts as a ligand for the cell surface-bound polypeptide receptors and stimulates proliferation, survival, and differentiation of the erythroid progenitor cells22-24. EPO is produced in mammalian cells and still generates over 12 billion US dollars a year. California-based Amgen, currently the largest biotechnology company in the world is the market leader in EPO (Epogen) production.

An example of a therapeutic protein used for the treatment of inflammatory disorders is etanercept (Enbrel), which is marketed by Amgen and Pfizer. It is a very successful therapeutic protein in terms of commercial success with sales of 3.5 billion US dollars in 2009, but above all in terms of therapeutic success. In contrast to the above described MAb Humira, Enbrel is an artificially engineered fusion protein that blocks Tumor Necrosis Factor (TNFα)25. TNFα is
one of the most common pro-inflammatory cytokines responsible for various inflammatory disorders, including rheumatoid arthritis\textsuperscript{18}.

As a final example, interferon beta-1b (Betaseron), produced by Schering/ Bayer is a very successful therapeutic protein for the treatment of multiple sclerosis\textsuperscript{26,27}, with sales of 4 billion US dollars in 2009. The exact mechanism of action of Betaseron is unknown, but immunomodulatory effects of Interferon beta-1b include the enhancement of suppressor T cell activity, reduction of pro-inflammatory cytokine production, down-regulation of antigen presentation, and inhibition of lymphocyte trafficking into the central nervous system.

1.2.3 Vaccines

Commercial interest in vaccines diminished in the 1980s and early 1990s because of the widespread use and effectiveness of antibiotics, and the low profit margins and liabilities associated with vaccine products. Several factors, such as new developments in modern biotechnology and immunology, and the danger of an influenza pandemic, have driven a recent spate of renewed interest in vaccines. Vaccines are also viewed as solutions in fighting poverty related diseases in the third world. Preventive vaccines are considered relatively cheap, in comparison to the costs of illness and treatment and the huge losses in lives.

Currently, the global vaccine market is fairly evenly divided between pediatric and adult vaccines, each account for about half of the total vaccine market, valued at 18 billion US dollars for 2009. Scares about diseases as avian influenza and new emerging infections like swine flu are stimuli for the vaccine market. In addition, the introduction of cancer and rotavirus vaccines has greatly expanded the vaccine market. Vaccines are also being developed against new targets including Alzheimer's, drug addiction, HIV, multiple sclerosis, tropical diseases, and autoimmune disorders such as diabetes, lupus erythematosus and arthritis.

The use of vaccines can either be prophylactic, to prevent the effects of a future infection, or therapeutic. Originally, two types of anti-viral vaccines were used: killed vaccines (preparations of the normal infectious virus that has been rendered non-pathogenic) and attenuated vaccines (altered to a non-pathogenic form). To reduce the risk of adverse reactions, most of the new vaccines under development are based on well-defined molecular immunogens. These ‘molecular vaccines’ include proteins, peptides, lipopeptides, plasmid DNA, and recombinant viruses based on viral vectors known to be safe in humans.

Human papillomavirus (HPV) is a sexually transmitted virus, and it is estimated that nearly 50% of sexually active people will contract it at some time in their lives. In most cases, the viruses are neutralized by the immune system over a period of two to three years. In 70% of all cases of cervical cancers, high-risk varieties of HPV are involved\textsuperscript{28}. HPV vaccines offer protection against the pathophysiological effects of HPV infection and HPV-infection-induced cervical cancer\textsuperscript{29}. Vaccination is the best solution to HPV infection and its effects, as no cure exists. The mostly used vaccine for HPV is Gardasil/Silgard from Merck \& Co., but also Cervarix from GlaxoSmithKline is being used. The FDA recommends vaccination before adolescence and potential sexual activity. In the Netherlands, a national vaccination campaign was started in 2009, to widely vaccinate twelve year old girls with Cervarix. The emergence was disappointing, only approximately 50% of the girls got vaccinated in 2010. Concerns about safety and effectiveness of the vaccination were discussed on the internet and this was thought to be the cause of the low emergence. This indicates how important proper and understandable information about vaccination is. This said, one has to take into account that only 70% of cases of cervical cancer can be prevented with this vaccination. In addition, one unknown property of the vaccines now being researched is the persistence of their protective effects. GlaxoSmithKline currently states that protection is warranted for at least 6 years and 5 months. In the Dutch journal ‘Medisch Contact’, certain physicians expressed their concerns of reduced alertness for cervical cancer when women consider themselves safe due to the
vaccination. This could cause detection of cervical cancer at later, more pathogenic stages. Again, this indicates that proper information can be of utmost importance for the success of certain biopharmaceuticals.

Above described Gardasil/Silgard and Cervarix are examples of preventive vaccines, which are designed to cure the condition (usually a virus) causing the cancer. Research in the vaccine field has also led to the development of therapeutic vaccines against cancers, which are designed to treat the active disease. The idea of using therapeutic vaccines against cancers is long-established, but hard to accomplish. Advances have been made over the past few years in developing cancer vaccines using whole-cells, proteins or peptides, plasmid DNA, and viral vectors with new adjuvants. The idea behind therapeutic vaccines is that ‘awakening’ the immune system to the presence of cancer by presenting it with antigens associated with tumor cells would lead to a sustained body wide search for similarly suspicious cells and also retaining a memory of the abnormal antigens. For instance, approved in 2010 by the FDA is sipuleucel-T (Provenge) by Dendreon, a whole-cell vaccine for the treatment of metastatic, hormone-resistant, prostate cancer (HRPC). The vaccine is a patient-specific vaccine, produced by incubating an individual’s own blood, enriched for dendritic cells and other antigen-presenting cells with a recombinant fusion protein composed of prostatic acid phosphatase (PAP) and granulocyte-macrophage colony stimulating factor (GM-CSF). Although not tumor specific, PAP is highly tissue specific and expressed in the majority of prostate tumors. In hormone-refractory prostate cancer patients treated with the vaccine, the median survival was significantly longer than placebo\textsuperscript{30,31}. Provenge is the first therapeutic cancer vaccine to demonstrate effectiveness in Phase III trials by prolonging life of patients who have advanced to the late stage of the disease.

1.3 Production vehicles of biopharmaceuticals

Biopharmaceutical production vehicles include microbial cell culture, yeast cell culture, animal cell culture, transgenic plants and transgenic animals. Choosing the most appropriate method for commercial production requires case-by-case analysis. A wide range of factors must be considered, including cost of production, market volume, the efficacy, safety and stability of the product, and whether the produced biopharmaceutical has the required biochemical or pharmacological properties.

1.3.1 Bacteria

The simplest means to produce biopharmaceuticals is in bacteria. In particular non-pathogenic bacteria such as \textit{Escherichia coli} grow fast and culture conditions are relatively simple and with low costs. In addition, bacteria are a rather robust organism, \textit{E. coli}’s genetics are well-characterized and there is an increasingly large number of cloning vectors and mutant host strains available. In all, this makes production of biopharmaceuticals much more economical than production of the same agents in more complex and vulnerable mammalian cells\textsuperscript{32}. There is, however, a huge drawback to bacterial production systems. Bacteria are incapable to provide proper post-translational modifications to the desired protein. Often, these post-translational modifications are essential for the therapeutic protein, in terms of biological activity, efficacy and half-life. Therefore, bacteria are only used in cases of relatively simple biopharmaceuticals (peptides, small proteins) where post-translational modifications are absent or unimportant for the biological activity. This is, for instance, the case with insulin that does not undergo glycosylation in its natural human form and is a small protein anyway. Therefore, insulin is still produced by bacteria (Ely Lilly). In addition, it remains one of the major challenges in bacteria to obtain the biopharmaceuticals in a soluble and bioactive form. Human proteins that are expressed at high levels in \textit{E. coli} frequently acquire an unnatural conformation owing to lack of proper folding and disulfide bridges. Moreover, high protein accumulation often results in
aggregation (inclusion bodies), thus, laborious renaturation steps are needed to recover the proteins in their biologically active form.

1.3.2 Yeast

Being a simple eukaryotic cell, yeast combines fast growth and easy culturing methods with an extracellular expression. However, its post-translational modifications are yeast-specific. For the production of therapeutic glycoproteins intended for use in humans, yeast has been less useful owing to their inability to modify proteins with human glycosylation structures. Yeast N-glycosylation is of the high-mannose type, which confers a short half-life in vivo and thereby compromises the efficacy of most therapeutic glycoproteins. Large research efforts have been made to introduce the human genetic machinery of post-translational modifications into yeast, but so far, this has not resulted in major breakthroughs. In addition, practical experiences show a high amount of product loss due to degradation of the target protein in the medium. Despite these limitations, *S. cerevisiae* is an established host organism for a number of recombinant proteins (for instance the urate oxidase rasburicase (Fasturtec) and the platelet-derived growth factor becaplermin (Regranex)), where the yeast glycosylation is not an issue in clinical practice.

1.3.3 Plants

For plants the same reasoning applies that growing a genetically modified plant in the field is cheaper than large bioreactor systems for mammalian cells. Although genetically engineered plants meet with a lot of resistance worldwide, predictions are that when such a plant would encompass a medicine, this public resistance would fade away. Unfortunately, post-translation modifications are plant-specific; complex-type N-glycans are structurally different in plants and mammals and therefore hardly useful for many biopharmaceuticals. Furthermore, purification of therapeutic agents from plants may encounter pesticides, herbicides as well as plant protein glycans that evoke allergic reactions. In addition, silencing of introduced transgenes has frequently been observed in plants, constituting a major commercial problem. Finally, proteolytic degradation of the biopharmaceuticals is another obstacle characteristic of plant-based systems. However, targeting signals can be used to intentionally retain recombinant proteins within distinct compartments of the cell to protect them from proteolytic degradation, preserve their integrity and to increase their accumulation levels.

Still, plants are considered as a promise, when certain issues can be resolved. As in yeast, attempts are being made to ‘humanize’ the plant post-translational machinery. Furthermore, when the biopharmaceutical is not purified from the plant, but when the plant is eaten, this may circumvent allergic reactions. Because humans are constantly exposed to plant glycoproteins in the diet, glycosylated plant-made biopharmaceuticals should be acceptable for oral administration. This scenario has been envisioned for the incorporation of vaccines in the plant genome, to generate edible oral vaccines. Particularly for the third world these edible vaccines would be a solution to many logistic problems surrounding the question how to get the vaccines to people in hardly accessible places.

1.3.4 Mammals

Production of biopharmaceuticals in animals is both very controversial and still in its infancy. Milk is presently the most mature system to produce recombinant proteins from transgenic organisms. Besides public concerns of transgenic animals, a drawback of production of biopharmaceuticals in animals may be the difficulty to separate the human proteins from their animal counterpart. Animal systems also suffer from long development timelines and possible contamination of purified proteins with animal viruses and prions. Moreover, some of the
recombinant proteins may be active and deleterious for transgenic animals. Hardly any therapeutic agent has reached the market that has been produced by a genetically modified animal. Pharming announced in 2010 that conestat alfa (Ruconest), a human C1-inhibitor produced in the milk of rabbits was approved for the treatment of angioedema attacks in patients with hereditary angioedema in all 27 EU countries plus Norway, Iceland and Liechtenstein. Dutch based Pharming has a long history as a biotech company with the goal to produce biopharmaceuticals in animals, for instance rabbits, goats or cows. One quote of Herman de Boer, founder of Pharming was that one transgenic flock of goats in a meadow would out-produce an entire, 100 million dollars worth production facility for mammalian cells. So far, this prediction has not become reality though.

1.3.5 Mammalian cells

Today, about 60-70% of all recombinant therapeutic proteins are produced in mammalian cells. The most widely used hosts for therapeutic protein production are Chinese Hamster Ovary (CHO) cell lines (mostly derivatives of the original CHO-K1 cells). Other cell lines used include cell lines derived from mouse myeloma (NS0 and Sp2/0-Ag14), baby hamster kidney (BHK) cells and human embryo kidney (HEK-293) cells. The advantages of mammalian cells are the proper protein folding and assembly and that proper post-translational modifications are achieved. For instance, the glycosylation machinery is highly conserved among mammalian cells and proper protein folding, efficacy, functionality or half-life of the biopharmaceutical heavily depends on the glycosylation characteristics of the protein. Therefore, biologically fully active biopharmaceuticals can be expected to be expressed in mammalian cells. Although there are marginal differences between human and non-human mammalian glycosylation, these differences in glycosylation potential do not appear to result in glycoproteins that are immunogenic. This is also apparent from the numerous biopharmaceuticals produced in mammalian cells that have been successfully used world-wide for many years. Note that not all mammalian cells are suitable for biopharmaceutical production. For instance, mouse cells generate proteins with N-acetylglucosamine (GlcNAc) residues that are highly immunogenic in humans, although there are exceptions (e.g. NS0).

To make a full-length antibody molecule, mammalian expression is the system of choice. Mammalian expression systems have been proven to generate safe and effective antibody molecules with serum half-lives equivalent to those observed for naturally occurring antibodies. HAHA (human antihuman antibody) responses owing to unusual carbohydrate structures have not been observed. For the production of (complex) therapeutic proteins with correct post-translational modifications being important for proper functioning, mammalian expression platforms are also preferred. Many proteins that are glycosylated in their natural form fail as therapeutics when not glycosylated. The most prominent example is EPO, which when originally produced in E. coli was fully active in vitro but not efficacious in vivo. The cause was an insufficient pharmacokinetic profile due to the absence of glycosylation. The mammalian cell systems have an additional advantage: recombinant proteins are secreted into the media in a natural form. Proteins expressed in E. coli, on the other hand, mostly accumulate within the cell as inclusion bodies in a highly denatured form, from which they have to be renatured during further manufacturing.

Unfortunately, when compared to E. coli, negative aspects of mammalian cells are that they grow relatively slow, are difficult to culture and they are very fragile. Also, expression levels of therapeutic proteins in mammalian cells are relatively low and instable. The productivity of mammalian cell culture processes is typically ~10-100 fold lower than what can be achieved using microbial host systems. Whereas the volume of microbial cells can correspond to 20-30% or more of the total culture volume, the cell biomass only occupies 3-4% in mammalian
cell cultures. This all adds up to the main disadvantage that the production of biopharmaceuticals in mammalian cells is vastly more expensive than production in *E. coli*.

Because most genes can be expressed in many different systems, it is essential to determine which system offers the most advantages for the production of active material at the lowest cost. From one safety point of view, it would theoretically be best to choose an expression platform that is phylogenetically most distant from humans since human pathogens cannot contaminate such systems. If glycosylation is required, then normally mammalian expression platforms are favourable, such as CHO cells. The first human therapeutic protein, produced in mammalian cells, was tissue plasminogen activator (tPA). It obtained approval by the FDA in 1986, hardly three decades ago. Hence it is safe to state that production of biopharmaceuticals in mammalian cells is still in its infancy. During these last three decades many improvements have been made in terms of cell culture conditions and the expression levels of the desired proteins. In particular the last aspect is also the scope of this PhD thesis.

# 2 Production of Biopharmaceuticals

## 2.1 Industrial production process

During the past three decades there has been one mammalian cell line that has been used for most produced biopharmaceutical proteins: immortalized CHO cells. Their suitability for large-scale and stable productions of biopharmaceuticals is related to their advantages of safety for use in humans, resemblance between glycan structure of their product with the natural human protein or MAb, ease of transfection, presence of a gene amplification system, ease of adaptation to growth in suspension and serum-free medium, and the ability to grow at high densities. Whereas these cells were in the beginning grown as adherent cultures, this has major disadvantages. Adherent cell culture processes are more demanding at large scales owing to the high surface to volume ratio needed to maximize cell densities. One solution that has been found was to grow adherent cells in large roller bottles, with cells grown attached to the inner wall of the vessels filled with medium. This is still the means by which EPO (Epogen) is produced by Amgen. However, most biopharmaceutical-producing mammalian cells are currently grown in suspension, allowing efficient upscaling processes.

Another disadvantage of adherent cell cultures is that the cells need serum in the culture medium to facilitate the attachment and spreading of cells. In the early days of cell cultivation, almost all culture media contained serum as a growth-promoting component. Serum was shown to have several essential functions in culture, for instance, it is a source of nutrients, hormones, growth factors and protease inhibitors. Besides these growth-promoting properties, serum has some major disadvantages. It is variable and undefined with respect to its chemical composition, there is variation between individual batches and it may contain substances that induce unwanted cell behaviour. As a result of their high protein content, the downstream processing of serum-containing culture media is more difficult and costly. Moreover, suspicions are that serum might contain contagious agents such as viruses and prions. Currently only serum is allowed that comes from countries that are proven to be free of mad cow disease, such as New Zealand. It is therefore commonplace these days to culture the mammalian cell cultures in serum-free suspension medium. This, unfortunately, creates another problem: due to growth-inducing components in serum, mammalian cells grow worse in serum-free medium. It took the industry many years to develop protocols to allow mammalian cells to grow reasonably well in serum-free suspension medium. Besides growing the cells, to create stable recombinant protein-expressing cell lines, transfected cells have to be selected and subcloned, which is more difficult with suspension cells. To overcome these problems caused by the absence of serum, balanced, chemically defined culture media have
been developed by several companies. The precise contents of these media are kept as well
protected and strict secrets. More general, the ability to adapt various cell types to suspension
culture and the use of polymeric additives to reduce shear damage have enabled the widespread
application of suspension cell culture.

Beside CHO cell lines (mostly derivatives of the original CHO-K1 cells) other mammalian
cell lines are also used: NS0 cells, Sp2/0-Ag14 cells, BHK cells, HEK-293 cells and human
retinal cells have gained regulatory approval for recombinant protein production. Still, however,
CHO cell lines are the standard when it comes to industrial production cell lines.

The most common procedure for the production of biopharmaceuticals in mammalian
cells is to establish stably transfected cell lines. This involves elaborate selection procedures
(see below). In another approach cell cultures are transfected transiently with DNA constructs
that encompass the gene of interest. In the resulting pools of cells relatively small amounts of
protein of interest are produced during a limited period of time, usually no more than a few
days in a bioreactor up to a 5-10 liter scale. Whereas this may sound as unproductive, these
small amounts of desired protein can be enough for initiation of quality studies or clinical
studies. For instance, for clinical phase I trials in which the safety of the biopharmaceutical in
question is studied, only small amounts suffice. Transient transfection in mammalian cell pools
is rapid and produces enough for these studies. Of course this is entirely different when it
comes to large-scale production of blockbuster MAbs such as Avastin. The production of such
proteins require the presence of large facilities which multiple, up to 20,000 liter bioreactors
that are solely dedicated to the production of this one protein.

What does a typical protocol for the production of biopharmaceuticals look like? Industrial
production processes are reviewed in Birch and Froud (1994), Kretzmer (2002) and
Rodrigues et al. (2009). As said, first the DNA construct encompassing the protein of interest
encoding gene has to be transferred to the cells. With suspension cells the currently most
commonly used method is by electroporation of the plasmid DNA into the cells. After that,
cells that have been transfected have to be selected, that is, non-transfected cells have to be
omitted from the cell culture. Selection of transfected cells (as also described below) is a major
topic in this PhD thesis and is described in Chapters 6, 7 and 8. Once individual clones that are
transfected have been selected, they are evaluated for their capacity to produce the
recombinant biopharmaceutical. However, not only highest producer cells are important. The
growth characteristics of the cell line are of importance as well. While a cell line may be able to
produce large amounts of the desired protein per cell, when these cells don’t grow, they are of
no use for further characterisation. In the end they need to be capable to grow in large-scale
bioreactors, up to as high as possible cell densities.

After these initial phases, multiple cell lines that produce the protein of interest are cultured
in small flasks and later in small bioreactors (1 to 5 liter scale). During this phase culture
conditions are tested and optimised for the particular cell line. Also the properties of the
protein of interest are tested. It is not, unfortunately, straightforward that each cell line will
produce a protein with identical properties. Hence the term ‘biosimilar’ is not invented for any
reason. The proteins that are produced at these small scales can be in enough amounts to allow
the initiation of clinical trials, certainly of phase I trials. Also during this phase the downstream
processing of the biopharmaceuticals is tested and optimised.
Although not of major interest to this PhD thesis, downstream processing is a major issue in the production of biopharmaceuticals. It involves the removal of undesirable agents in the cell culture, as well as undesirable derivatives of the product itself such as degradation products and aggregates. The optimal formulation of the end-product needs to be isolated. It also tests the ability of the purification steps to remove a range of viruses. In all it can be said that downstream processing constitutes the main factor in the ‘cost of goods’ of biopharmaceuticals.

Finally, when all conditions have been met satisfactorily, the scaling-up process of the cell culture takes place. Most biomanufacturing platforms involve either a batch or a fed-batch protocol, but other production options are available. A batch culture is the simplest process that involves no supplementation during a production run. Cultures are inoculated, followed for cell viability, and harvested at the appropriate time. Culture longevity is limited by exhaustion of nutrients and build-up of waste products. A typical fed-batch protocol involves adding concentrated supplemented components to a culture after inoculation. The concentrates are designed to supply needed nutrients that are consumed by the culture. As explained, after harvesting, downstream processing is involved in the recovery and purification of the biopharmaceutical. It may be clear from this description that it takes many (costly) years to come from a DNA construct to a biopharmaceutical that ends in a for sale glass bottle.

2.2 Selection of biopharmaceutical-producing mammalian cell lines

Above, we described the industrial production of biopharmaceuticals. To create biopharmaceutical-producing cell lines, the mammalian cells are stably transfected. How do you identify a cell line that contains the desired gene of interest and that produces the corresponding protein of interest at a high level? A wide variety of systems for selecting transfected cells exist, but most include the use of resistance to antibiotics such as neomycin, Zeocin, hygromycin or puromycin. Cells will die unless they harbor a gene that encodes a protein that is able to neutralize such antibiotics. Therefore, such a gene is always co-transfected in some form with the gene of interest, encoding the desired biopharmaceutical. The common procedure is that upon transfection of the gene of interest, along with the antibiotics resistance gene, the cells are subjected to the antibiotic in question. Dependent on the stringency of the selection system (see below) and the type of antibiotic used, the cells that are not transfected will be killed rapidly by the antibiotic. Only cells that have been transfected and that express a functional antibiotic resistance protein will survive and outgrow to a stable cell colony and eventually cell lines. It is presumed that along with the expression of the functional antibiotics resistance protein, also a functional protein of interest will be expressed. This is, unfortunately, not automatically the case.

There are a few important considerations to take into account when choosing a selection agent, for instance overall toxicity. Use of penicillin and other beta lactam antibiotics in media is not allowed for production of therapeutics and the use of antibiotics in general is discouraged. This has clearly to do with the difficulties to purify the antibiotics from the culture medium during downstream processing. Another important aspect to consider is the stringency of the selection system. For instance, the use of neomycin as selection agent is popular, but the neomycin resistance protein is an enzyme. This implies that only small amounts of this enzyme have to be expressed in the cell for the neutralization of many neomycin molecules. This in turn implies that the expression levels of this protein don’t need to be high in order for the cell to survive. Unfortunately, this is often accompanied by the protein of interest being expressed at a low expression level. As a result, selection with neomycin as selection agent always involves the induction of many transfected colonies (in the thousands) that almost invariably display low protein expression levels. Large-scale screenings
have to be performed to identify the rare colony that by chance displays high protein-expression levels.

One means to improve the selection stringency is the use of an antibiotic resistance protein that is not an enzyme. Only a few are available, and the Zeocin resistance protein is one of them. Zeocin belongs to the structurally related group of bleomycin/pleomycin type antibiotics. The cytotoxic effects of Zeocin result from the ability to cause fragmentation of DNA. Zeocin binds to DNA through its amino-terminal peptide, and the activated complex generates free radicals that are responsible for scission of the DNA chain. To neutralize this damaging effect of the agent, a gene that encodes the Zeocin resistance protein is required that binds and prevents the selection agent to cleave DNA\textsuperscript{62,63}, which is co-transfected with the gene of interest. The fact that this binding happens in a one to one ratio warrants that high Zeocin resistance protein expression levels are needed to neutralize enough Zeocin agent to prevent that this toxin will kill the cell. We routinely use Zeocin as selection agent in order to obtain transfected colonies that also display reasonable to high protein-expression levels.

Beyond the use of Zeocin as high stringency selection agent, our group also developed a very stringent selection system, based on the modification of the selection marker protein itself. We called this system STAR-Select and it utilizes attenuated translation initiation codons to make the translation of the selection marker protein less efficient\textsuperscript{64,65}. By creating a bicistronic mRNA, this system requires that high mRNA levels need to be produced in order to translate sufficient selection marker protein to allow the cell to survive. The concomitant, positive side effect of a bicistronic messenger is that the protein of interest that is translated with optimal efficiency is also expressed at a high level. This very stringent STAR-Select system allows the creation of only few colonies that almost invariably display very high protein expression levels. This is achieved in a very short time, during a first selection period. This is in contrast with the dfhr-mediated gene amplification protocol that can also reach high protein expression levels, but that requires multiple rounds of gene amplification and very long time lines (described below).

There are, unfortunately, drawbacks to the STAR-Select system. One is that in practice only a single translation initiation codon is available for use in this system: TTG, instead of ATG. Other sub-optimal translation initiation codons are either much too stringent, and no cells survive, or far less stringent and no major gain in high protein expression levels are achieved. We, therefore, sought means to overcome this problem and created a novel selection system that is more versatile than the STAR-Select system and that almost allows titration of the optimal selection stringency in a given cell line. Also this system plays with translation initiation frequencies. It is described in Chapter 5 of this PhD thesis.

A drawback of any selection system is that there is a need to keep antibiotics in the culture medium in order to obtain stability of gene expression over prolonged periods of time. When the antibiotics are omitted, most often the expression levels decline sharply within a few weeks. However, keeping antibiotics in the culture medium is not favored because of high costs and the need to omit such contamination during the downstream process development. One way to circumvent this situation is the use of so-called metabolic selection markers. That is, use a gene as selection marker that encodes for an essential step in the cell metabolism. Examples can be genes that encode proteins vital for the synthesis of essential amino acids. Normally, such amino acids need to be present in the culture medium for the cell to survive. However, this need can also be supplemented by adding this gene (that lacks from the cell in question) to the expression cassette and omit the essential amino acid from the culture medium. The co-transfected gene will produce the wanted protein (enzyme) and the cell will thus be able to survive, without the essential amino acid in the culture medium. Almost always there remains the need for the inclusion of a precursor agent in the culture medium that is needed for the synthesis of the essential amino acid. The usefulness of the system depends on the nature of
this precursor. It must be easily available commercially, it must enter the cell, and by preference, the working concentration must be non-toxic. If these requirements are fulfilled, this may result in the establishment of a selection system that does require modified, designed culture medium, but that is devoid of toxic agents that have to be removed from the final product.

In this PhD thesis we have elaborately sought for such metabolic selection markers. In Chapters 6, 7 and 8 we describe the genes that are involved in respectively the biosynthesis of L-threonine, hypoxanthine/thymidine/glycine and L-cysteine. We test whether these genes can be used for the direct selection of mammalian cells. As shown in Chapter 8, in particular the gene involved in the synthesis of L-cysteine seems suitable for this purpose.

2.3 Improvements in biopharmaceutical production and biopharmaceutical characteristics

High clinical dose requirements of biopharmaceuticals and the desire to minimize the ‘cost of goods’ are two major issues that demand for large-scale, efficient and cost effective biopharmaceutical production. Even modest increases in product yield can meet these demands. Alternatively, improving characteristics of biopharmaceuticals that allow for reduced administration is another way to tackle the high demand issue. Improvements of biopharmaceutical production at distinct levels are discussed. Figure 2 shows a schematic diagram of the biopharmaceutical production process.

**Vector optimization:**
- Strong cellular promoter (Ch.2)
- RPL32 upstream of primary promoter (Ch.3)
- DNA elements that augment gene expression (Ch.4)
- Adjustable stringent selection system (Ch.5)
- Metabolic selection markers (Ch.6,7,8)

**Selection:**
- Adjustable stringent selection system (Ch.5)
- Metabolic selection markers (Ch.6,7,8)

**Screening ➔ high expression levels due to:**
- Strong cellular promoter (Ch.2)
- RPL32 upstream of primary promoter (Ch.3)
- DNA elements that augment gene expression (Ch.4)
- Adjustable stringent selection system (Ch.5)

**Stability of those high protein levels:**
- Metabolic selection markers (Ch.6,7,8)

**Figure 2.** Schematic representation of a biopharmaceutical production process. First steps in the production process are molecular cloning of the gene of interest in an expression vector, which also contains a selection marker. Next, the plasmid DNA containing the gene of interest and selection marker is delivered into the cells by transfection. Following DNA transfer, cells are subjected to selective conditions to recover those that have stably integrated the exogenous genes into a chromosome. Clonal cell lines are isolated and evaluated for biopharmaceutical expression and other desired characteristics. The improvements in the production process that we present in this thesis and the respective chapters in which they are described are indicated in the text boxes. Next steps in the production process include up-scaling of biopharmaceutical-producing cells and downstream processing. Adapted from Mellstedt et al., 2008.
representation of a typical production process and the improvements in protein production that we present in this PhD thesis are displayed.

2.3.1 Gene amplification
The current main approach to obtain higher therapeutic protein expression levels is through gene amplification. This is achieved through the involvement of the Dhfr (dihydrofolate reductase) gene. DHFR is an enzyme that converts dihydrofolate into tetrahydrofolate, a methyl group shuttle required for the de novo synthesis of purines (hypoxanthine), thymidyl acid (thymidine), and the amino acid glycine. The commonly used CHO-DG44 production cells lack the Dhfr gene and these cells therefore need glycine, hypoxanthine and thymidine in the culture medium to survive. When the Dhfr gene is placed on the expression cassette that also harbors the gene of interest, the cell can convert the non-toxic substance dihydrofolate into tetrahydrofolate, bypassing the need for exogenous glycine, hypoxanthine and thymidine. Generally, selection of transfected cells is accomplished by simultaneously selecting with an antibiotic such as neomycin. Once stable cell lines have been established, the activity of the DHFR enzyme is inhibited by addition of the specific inhibitor methotrexate, a toxic substance. The cells react to this by amplifying the entire gene locus, including the gene of interest. Since this results in many more copies of the gene of interest (up to 10- thousands of copies after multiple amplification rounds), it can be expected that also the expression levels of the protein of interest are elevated. While this is true, the methodology also has its disadvantages.

Usually multiple rounds of gene amplification are needed to achieve adequate protein expression levels. That is, after a first round of methotrexate-induced gene amplification, a cell line is chosen that displays high protein expression levels. Subsequently, such a cell line is again subjected to higher doses of methotrexate to induce even more gene amplification and this is repeated until sufficient transgene expression is obtained. Generally, amplification of the Dhfr gene requires 12 weeks per cycle and it can take up to 5 cycles to obtain a clone with high enough expression and good growth characteristics for today’s biopharmaceutical products. The identification of an amplified colony that displays higher protein expression levels is a tedious process, because the extend of amplification ranges widely between clones. Often hundreds of cell lines have to be characterized. Also, while the end result may be a cell line that produces high protein levels, the cell is often genetically very unstable due to the entire amplification procedure. Also, the continuous presence of toxic methotrexate is required in order to achieve stability of protein expression over prolonged periods of time. When methotrexate is removed from the cell culture, the cell tends to return to a previous state, with a gradual loss of the gene copy number. Obviously, this is accompanied by lower protein expression levels, not a desired outcome. In Chapter 7, we present a means to use DHFR to maintain stable protein expression, without the need for methotrexate-induced gene amplification.

2.3.2 Optimization of culture conditions
Above it has already been mentioned that altering culture conditions of the cell culture may profoundly influence the behavior of the cell culture and the ability to secrete high amounts of biopharmaceuticals. During the last few decades considerable progress has been made in improving the productivity of mammalian cell processes as the result of systematic production process optimization, including optimization of culture media. One significant development (as discussed above) is the change towards serum free suspension cell cultures. Many commercial media for cell culture are now available from a few leading suppliers. Given the importance of cell culture medium for the eventual outcome in terms of protein production, leading recombinant protein manufacturers also invest heavily in optimization of their own media
formulations, dedicated towards their own products and (proprietary) cell lines. Culture medium development has to be performed on an individual basis, for each process, each particular biopharmaceutical and cell line, that is, even dedicated to one cell line producing one specific biopharmaceutical. Taking this even further, several different media formulations are often necessary for a single manufacturing process, each one of them designed for a specific phase. This can be media for rapid growth that have formulations different from those of production media (a batch (6-8 days) or extended batch (10-21 days) process). One of the major advantages of defined culture media is that they are inherently less expensive. This is due to the fact that expensive serum is avoided and that downstream processing is much more straightforward as there are fewer contaminants (from the serum) to monitor and remove. Also, chemically defined culture media have been developed to dramatically improve the viability of cells when grown to high cell density, and thus delay the onset of apoptosis.

Process development is an empirical technology and science that involves the modification of many factors that are inherently linked to each other. Influencing one parameter is likely to modify another. Factors such as temperature, pH of the medium, carbon metabolism, the accumulation of (toxic) side products are all intertwined in a very complex system that eventually determine how well cells will produce the valuable biopharmaceutical. Environmental control strategies continue to be useful for manipulating those processes.

2.3.3 Translation and secretion

Whereas increase in the transcription rate of a heterologous gene is aiming to result in a higher amount of a secreted biopharmaceutical, this is by no means certain. The increased mRNA levels have to be faithfully followed by an at least similar efficiency of translation of the mRNA and subsequently by at least as efficient folding (and assembly) and secretion of the protein. Neither one of these aspects is for certain and can in fact become a novel bottle neck in the chain of events. To optimize the translation of a protein of interest, often optimized translation codons are introduced to replace codons that are less efficiently translated; minor tRNA species are avoided.

All secreted mammalian proteins are post- or co-translationally translocated into the lumen of the endoplasmic reticulum (ER) which contains the chaperone and enzyme proteome required to catalyze reactions by which proteins are modified, folded and assembled to attain their functional conformation essential for transport to the cis-Golgi compartment. Various attempts to engineer ER resident chaperone systems to boost the folding capacity of the ER of mammalian host cells resulted in small if any increases in cell specific productivities. One possible explanation for these results is that over expression of one chaperone may not be sufficient to increase the folding capacity of the whole folding machinery in the ER and that coordinated elevation of expression of several chaperones and foldases may be necessary to obtain higher and more reproducible increases in cell specific productivities. Multiple efforts have been made to identify and employ upstream effectors that are at the hierarchical top of the folding and secretion machinery. Amongst these are proteins such as XBP-1 (X-Box-binding protein 1) and ATF6 (Activating transcription factor 6), which are both involved in the unfolded protein response (UPR), which is activated to maintain the balance between the folding capacity of the ER and the folding demand imposed on this organelle. The influence of over expression of such regulatory proteins on transgenic protein expression levels in production cell lines have been met with variable success; opposing claims have been reported. This may be due to the fact that in mammalian cells, engineering of UPR signaling is complicated by the existence of several signal transduction pathways, extensive crosstalk between these pathways, and simultaneous activation of responses beneficial and detrimental to recombinant protein production. An important aspect of UPR engineering will
be to selectively activate those responses beneficial to protein production, while responses detrimental to protein production are not elevated or even abolished.

2.3.4 Anti-apoptosis and proliferation control

Once a good-producing cell line that also grows well is transferred to a (large-scale) bioreactor, the next problem arises. For typical fed-batch production processes, the ideal combination is a rapid accumulation of productive cellular biomass maintained at high viable cell concentration for as long as possible. When the cells have grown to a high cell concentration, it is vital that they stay alive and remain productive as secreting cells for generally 60 generations, depending on the length of the production run. This is, however, often not the case. Once a critical cell density has been obtained, cells may inherently follow a course towards apoptosis, controlled cell death. In fact, most cells die of apoptosis in batch cell cultures. Apoptosis in such cell cultures is hard to avoid, since apoptosis almost automatically follows when cells are put under extreme conditions and too much stress. And a production cell line that is pushed towards high secretion of a heterologous protein, under high density conditions with possible nutrient depletion, is inherently stressful for the cell. Since apoptosis is such a negative aspect for the production procedure of biopharmaceuticals, many strategies have been developed to limit cell death and to prolong culture viability.

In principle there are two complementing strategies to prevent or inhibit apoptosis: the cellular environment can be manipulated through media supplementation or the intracellular environment can be modified by genetic engineering. For instance, nutrient limitation can induce apoptosis; monitoring and adding nutrients is therefore of great importance, facilitating reduction of apoptosis and therefore increased accumulation of cellular biomass in vitro.

Various genetic engineering strategies to directly maintain or extend viable cell density have concentrated on inhibition of apoptotic cell death. Diverse approaches have been tested, most commonly expression of the survival (anti-apoptotic) proteins bcl-2 or bcl-xL, but also caspase inhibition, expression of the molecular chaperone HSP70 and RNAi suppression of pro-apoptotic proteins. However, the results of over expression of anti-apoptotic genes are contradictory with respect to productivity and a retarding effect of cellular proliferation has also been reported. There are as yet no reports describing successful application of this approach to improve biopharmaceutical yield from optimized industrial production systems.

Even if apoptosis is prevented, high density growing cells still will stop dividing. An accompanying phenomenon is that once they reach this stage, they also start to secrete more heterologous protein, a welcome side effect. Apparently, there is an inverse relationship between productivity and cell growth. This has been used to increase the specific productivity of cells. Over expression of p27, a regulatory protein that acts at the cell cycle G1/S phase transition, almost instantaneously leads to arrest of cell growth. Concomitantly, a 15-fold increase in the specific productivity (productivity per cell per day) was observed. How useful this finding may be in terms of industrial use remains to be seen, but the studies indicate how interference in a fundamental cellular process can result in a profound change in the production of a biopharmaceutical.

2.3.5 Glycosylation

Although beyond the scope of this PhD thesis, post-translational modifications of which glycosylation of proteins is one, is a very important aspect in the production of biopharmaceuticals. The ability of mammalian cells to provide proper glycosylation to biopharmaceuticals is the main reason to use ‘difficult’ mammalian cells as production platform in the first place. Whereas there are two forms of glycosylation of proteins, O- and N-glycosylation, only the latter is relevant for all recombinant human proteins that have been
reproduced to date. N-linked glycosylation involves the addition of sugar groups (glycans) to the amino acid asparagine within the polypeptide chain. Despite the fact that N-linked glycosylation involves more than a hundred genes, it is highly conserved across different mammalian species. The pattern of protein glycosylation is dependent on the expression of various glycosyltransferase and glucosidase enzymes that are present in the ER and Golgi apparatus of the cell. In general, pools of different glycoforms with varying glycan structures attached to a single peptide backbone with a known amino acid sequence are formed. For the production of recombinant proteins as a biopharmaceutical it is essential to ensure that a consistent glycosylation profile is maintained between batches. This may not be so easy to control given that the extent of glycosylation may decrease over time in a batch culture. Inconsistencies are likely to be due to the depletion of nutrients, particularly glucose or glutamine, which have been shown to limit the glycosylation process. Therefore, monitoring and adding nutrients is of great importance.

Why is the proper glycosylation of such importance? For instance, proper protein folding, efficacy, functionality or half-life of the biopharmaceutical heavily depends on the glycosylation characteristics of the protein. An example is antibody-dependent cellular cytotoxicity (ADCC), an important mechanism of action of therapeutic MAbs. Deletion of the FUT8 gene in CHO-DG44 cells results in cells that lack alpha-1,6-fucosyltransferase activity and cannot synthesise fucosylated antibodies. It has been shown that the ADCC of the resulting antibody was increased 100-fold compared to the fucosylated form. Also, over expression of N-acetylglucosaminyltransferase III in CHO cells increased the proportion of bisecting GlcNAc residues, increasing the ADCC substantially compared to the parental molecule. Therefore, oligosaccharide engineering has become an important research area for increasing antibody potency.

As a final example, many of the next-generation variants of the therapeutic EPO protein have extended glycosylation patterns. For instance darbepoetin alfa (Aranesp), a hyperglycosylated EPO, contains additional N-linked oligosaccharide chains that are accomplished by substitutions at five positions along the 165-amino-acid backbone without altering the tertiary structure. Increasing the sialic acid-containing carbohydrate content beyond the maximum found in natural EPO leads to a molecule with a longer circulating half-life and thereby an increased in vivo potency. This allows for less frequent administration, which is pleasant for the patient and also beneficial for the health care budget.

2.3.6 Improving transcription levels of the genes of interest

Over the years lots of attention has been given to improve the protein expression levels in mammalian cells through refinements of vector construction. A common problem in transfected mammalian cells is that expression levels of the gene of interest are simply low. Transcription in mammalian cells involves promoters, enhancers, introns, polyadenylation signals and so on. All of these aspects have been modified to obtain better gene expression levels. For instance, promoters and enhancers can be very tissue specific and it is not clear what the optimal promoter-enhancer combinations are for CHO cells. Even worse, the culture conditions of these cells, such as serum free suspension medium may have a huge impact on the outcome of such studies.

The search for optimal promoter-enhancer combinations has resulted in, for instance, the CAG promoter. This is a combination of the CMV-IE enhancer and a modified chicken β-actin promoter, which is able to confer high activity in cultured cells. An expression vector containing transcription regulatory sequences from the Chinese Hamster EF1α gene has also been developed for high-level expression of proteins in mammalian cells.

Beside the identification of novel promoter enhancer combinations, others have attempted to over-express genes of interest with customized transcription factors. These are derived from
randomised zinc finger protein (ZFP) libraries. Identified ZFPs can be attached to a functional domain (transcription inducer) to generate a ZFP transcription factor (ZFP-TF). The engineered ZFP-TF can be used to upregulate expression levels in a given cell type and with a specific gene of interest. The USA-based company Sangamo follows a patented procedure to engineer ZFP-TFs with a given target promoter to obtain higher protein expression levels. Given the number of licenses they sold to large companies, including Genentech, this appears to be a fruitful approach. Drawback of these customized transcription factors is that besides the protein of interest, another protein (the ZFP-TF) has to be stably expressed in order to warrant high-level expression.

It can still be said though, that despite these efforts the most commonly used promoter in the biotechnology industry remains the CMV promoter, derived from the human cytomegalovirus immediate-early gene. Since the strength of the CMV promoter can vary considerably in different cell types and physiological conditions, we decided to compare a range of potent promoters in CHO cells. This study is presented as Chapter 2 in this PhD thesis.

We have given a novel twist to this story by the discovery that placing one specific promoter, a human ribosomal promoter (RPL32), upstream of another promoter strongly elevates the potential to induce high-protein producing cell lines. We describe this in Chapter 3 of this PhD thesis.

2.3.7 Strategies to counter transcriptional silencing

In the genome genes reside as chromatin, the complex structure of DNA and associated proteins, in particular the histones. At the molecular level, the DNA double helix is wrapped around histones and this constitutes the basic building block of chromatin: the nucleosome, repetitive histone octamer units typically enfolded by 147 base pairs of DNA. Acetylation, methylation, phosphorylation and ubiquitination of lysines and serines of amino-terminal histone tails alter the extent to which DNA is wrapped around histones and thereby have enormous impact on whether a gene is active or inactive.

When a heterologous gene, in our case the gene of interest, randomly integrates into the genome of the host, it may integrate in repressed or ‘closed’ (hetero)chromatin, resulting in gene silencing. In contrast, only when integrating in ‘open’ (eu)chromatin in which genes are in potency transcriptionally active, the transgene might also be transcriptionally active. Thus, efficient expression of the transgene is highly dependent on the place of its integration. Since a large proportion of the genome resides in the form of heterochromatin, the chance that a transgene integrates in or close to heterochromatin and consequently is silenced or repressed is also high. Even if the transgene has integrated into an active region of the genome and is expressed at high enough levels, it is common that the expression levels drop after several rounds of cell division. This gene silencing is thought to be associated with epigenetic factors resulting in the condensation and transcriptional inactivation of the chromatin at the site of integration. Importantly, once genes obtain a repressed heterochromatic state, this status is faithfully transmitted to daughter cells. A number of strategies has been developed that aim to overcome the effects of random integration and subsequent transcriptional silencing of the integrated transgene.

1. Targeting transgene integration to known transcriptionally active regions of the genome. This strategy includes the use of Cre-lox recombination. This involves targeting of gene expression cassettes that are flanked by LoxP sites to a cell line that has been designed to contain corresponding LoxP response sites in a well-defined chromosomal location. Upon transfection of the expression cassettes, a recombinase must be activated in order to integrate the cassette in the pre-destined location.
2. Creation of artificial chromosomes. An entire (small) chromosome has been ‘built’ that does not contain or is not prone to heterochromatin formation. Insertion of the transgenic gene expression cassette into such an artificial chromosome will, therefore, result in a continuous ‘open’ and active chromatin configuration of the expression cassette. The challenge with this system is to choose DNA sequences to build the 60 Mb artificial chromosome that are not prone to gene silencing and remain episomal during several rounds of cell division. Once this has been accomplished, the artificial chromosome that harbors the expression cassette must be isolated and transferred to a cell line. Obviously, it requires very specialized personnel to physically isolate the engineered chromosome and inject them manually into the nuclei of cells. It is thought that from then on, the artificial chromosome will stably propagate in the new production cell line\(^{135-137}\).

3. Various \textit{cis}-acting gene-regulatory elements have been identified and employed to enhance transgene expression. Most well-known are insulators, which are DNA sequences naturally occurring in the genome that block the action of enhancers and prevent repression from heterochromatin without a regulatory activity of its own\(^{138-140}\). Furthermore, there are so-called ubiquitous chromatin opening elements (UCOEs) that contain a set of divergent promoters and a CpG island\(^{141-142}\). These regulatory elements appear to confer a dominant chromatin opening function and thereby resist transgene silencing. Scaffold or matrix attachment region (S/MAR) elements bind the so-called nuclear matrix and are thought to be involved in creating higher order structuring of chromatin in topologically independent loops\(^{143-145}\). Finally, in our group we identified stimulatory and anti-repressor elements (STAR) elements that are thought to block chromatin-associated gene repression\(^{146}\).

Flanking transgenes with any of these elements aims to reduce the effects of heterochromatin and thereby allow stable and high expression of the transgene. In a comparative study, we showed that in the context of a very stringent selection system, in particular MAR and STAR elements were very effective in elevating (trans)gene expression in stably transfected cell lines\(^{147}\). In particular when it comes to the \textit{cis}-acting regulatory elements, most of these approaches have a strongly pre-meditated background. They heavily rest on theoretical insight in how the eukaryotic genome works. In Chapter 4 of this PhD thesis we describe a methodology that, without a pre-mediated vision, enabled us to isolate genomic sequences that operate at least as well as STAR elements.

3. **Outline of the Thesis**

In this chapter we introduced biopharmaceuticals, an important class of modern medicines. For the production of glycoproteins, mammalian cells, particularly CHO cells, are the host of choice. Unfortunately, low protein production levels are often observed in mammalian cells. Long time-lines to isolate appropriate clonal cell lines and instability of expression over time are also major problems that still hamper the efficient industrial production of biopharmaceuticals. Improvements in product yield are needed to meet the issues of high demand and the desire to reduce ‘cost of goods’. Improvements at distinct levels can increase product yield. In Chapters 2 and 3 we describe the potential of cellular promoters to induce a high number of mammalian cell lines in the context of a stringent selection system. Whereas Chapter 2 describes the behavior of a number of cellular promoters themselves, in Chapter 3 we combine these promoters with a second promoter that normally drives the expression cassette. We show that of 11 tested promoters, only the promoter of the large ribosomal protein RPL32 is able to raise the number of stable colonies, when the promoter is placed immediately upstream of for instance the CMV or \(\beta\)-actin promoter. The cell lines that result
from this induction display very high protein expression levels. In Chapter 4 we describe a novel screening method to identify genomic sequences that are able to induce large numbers of high-expressing colonies, in the context of the very stringent STAR-select selection system.

A novel selection system that almost allows titration of optimal selection stringencies is described in Chapter 5. In Chapters 6, 7 and 8 we describe the genes that are involved in respectively the biosynthesis of L-threonine, tetrahydrofolate and L-cysteine. We test whether these genes can be used as metabolic markers for the direct selection of mammalian cells. As shown in Chapter 8, in particular the gene involved in the synthesis of L-cysteine seems suitable for this purpose. In Chapter 9, the presented results are evaluated and discussed and future directions for the described research are suggested.

REFERENCES


61. Center for Biologics Evaluation & Research (CBER). Points to consider in characterisation of cell lines used to produce biologicals (2001).


