Technical Report

Recombinant therapeutic proteins: Production platforms and challenges

Theo Dingermann
Institute for Pharmaceutical Biology, Goethe University, Frankfurt/Main and Center for Drug Research, Development, and Safety (ZAFES), Frankfurt/Main, Germany

Since the approval of insulin in 1982, more than 120 recombinant drug substances have been approved and become available as extremely valuable therapeutic options. Exact copying of the most common human form is no longer a value per se, as challenges, primarily related to the pharmacokinetics of artificial recombinant drugs, can be overcome by diverging from the original. However, relatively minor changes in manufacturing or packaging may impact safety of therapeutic proteins. A major achievement is the development of recombinant proteins capable of entering a cell. Such drugs open up completely new opportunities by targeting intracellular mechanisms or by substituting intracellularly operating enzymes. Concerns that protein variants would cause an intolerable immune response turned out to be exaggerated. Although most recombinant drugs provoke some immune response, they are still well tolerated. This knowledge might result in a change in attitude towards antibody formation, i.e., neutralizing antibody activity (in vitro) may be overcome by dosing consistently on the basis of antibody titers and not only on body weight. As with other drugs, efficacy and safety of therapeutic proteins have to be demonstrated in clinical studies, and superiority over available products has to be proven instead of just claimed.

Keywords: Biotechnologicals · Glycosylation · Post expression modification · Recombinant proteins · Safety

Introduction

Human proteins produced through genetic engineering rather than isolated from tissue samples have become an important category within therapeutic medicines. Progress in production technology, as well as in pharmacological understanding, will allow continued development of human proteins as an important therapeutic option in a variety of human diseases.

From a clinical perspective, therapeutic proteins provide essential therapies in various life-threatening disease including diabetes (insulin), end-stage renal disease (erythropoietin), viral hepatitis [interferon (IFN)], cancer-treatment related neutropenia (G-CSF), clotting disorders (Factor VII, VIII, IX), and inborn errors of metabolism (lysosomal enzymes).

From a biochemical viewpoint, there is an interesting diversity within these proteins in their primary protein structure (chain of amino acids) compared to the naturally occurring human form(s) and their post-translational modifications (including the natural process of glycosylation, the addition of sugar groups to specific amino acids, and the artificial process of PEGylation, the addition of PEG groups to amino acids). Initially, biochemical differences between the therapeutic proteins and their natural counterparts often reflected either technical or biological limitations. Such deviations, however, do not automatically have an actual im-
pression control of genetic information is highly have to be adopted to the target organism, since ex-DNA-encoded information into RNA and proteins genes which control the translation process of the identical among different species. Only those re-building blocks of DNA, and the rules for encoding foreign protein. This was possible because the successfully introduced into a bacterium, resulting in copied it into a bacterial plasmid, which they suc-

pressed. They lifted human DNA and were the first to perform a type of experiment, initi-

vating what was later coined recombinant DNA technology. They lifted human DNA and

quired for producing insulin and human growth hormone. But soon it was realized that E. coli had one serious shortcoming: it was unable to modify proteins by glycosylation. This was not a problem for producing insulin and human growth hormone since these proteins do not undergo glycosylation in their natural human form. Problems were, however, anticipated for many other recombinant proteins with therapeutic value, since most of the human extracellular proteins (and even some intracellular proteins) are naturally modified with sugar chains.

Glycosylation of proteins is a highly complex post-translational modification process taking place in the endoplasmic reticulum and Golgi apparatus and involving more than a hundred different proteins (and genes). This glycosylation machinery is absent in E. coli, present but different from mammalian cells in the yeast Saccharomyces cerevisiae, but highly conserved among mammalian cells (e.g., human, Chinese hamster).
It came as a welcome surprise that even some therapeutic proteins like the IFN-α, -β and -γ, interleukin-2, tumor necrosis factor-α and others, which are glycosylated in their natural human forms, turned out to be effective medicines, even though they were not glycosylated after expression in E. coli.

However, other proteins failed as therapeutics when not glycosylated. The most prominent in this group was erythropoietin, which when originally produced in E. coli was fully active in vitro but not efficacious in vivo [3]. The cause was an insufficient pharmacokinetic profile due to the absence of glycosylation, and not a primary pharmacodynamic failure as receptor binding itself is not dependent on sugar modifications. Whereas there are two forms of glycosylation of proteins, O- and N-glycosylation, only the latter is relevant for all 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.

Although the yeast S. cerevisiae expresses N-glycosylation sites encoded by human genes, its glycosylation pattern is so different from that of humans [4] that concerns about whether such an “artificial” glycosylation of recombinant products would induce a severe immune response in humans were raised. Nevertheless, S. cerevisiae is in the meantime an established host organism for a number of recombinant proteins (human insulin, glucagon, the hirudin analogues desirudin and lepirudin, the urate oxidase rasburicase, and the platelet-derived growth factor becaplermin), where the yeast glycosylation is not an issue in clinical practice.

Use of three mammalian cell lines, derived from Chinese hamster ovary (CHO) cells, baby hamster cells (BHK) cells, or human fibrosarcoma cells, elegantly provided the necessary glycosylation in the production of many other therapeutically intended naturally glycosylated proteins. Despite the fact that N-linked glycosylation involves more than a hundred genes, it is highly conserved across different mammalian species. Marginal differences between human and non-human mammalian glycosylation include the absence of N-glycolyl neuraminic acid (NeuGc), a sialic acid, in human proteins. Humans lack the enzyme to modify N-acetyl neuraminic acid (NeuAc), the most predominant form of sialic acid in N-linked glycosylation in all mammalian proteins, into NeuGc. In all CHO cell-derived therapeutic proteins, NeuGc represents a minor fraction (ranging from less than one to several percent) of sialic acid contents. Whereas antibodies to NeuGc have been reported in humans, possibly triggered by food intake, antibodies against NeuGc have never been documented in any of the genetically altered CHO cell-derived therapeutic proteins, despite the fact that they have been used in millions of patients by now. Further support of the lack of clinical relevance is provided by the presence of complete cross-reactivity of antibodies triggered by the exposure to therapeutic proteins produced in a human fibrosarcoma or CHO cell platforms.
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, generally with limited although sufficient yields.

For a long period of time, basically all recombinant proteins that were approved as drugs were produced in just four cell systems: E. coli, S. cerevisiae, CHO and BHK cells. Frequently, a particular therapeutic protein was available from two different expression systems. Insulin and human growth hormone, for example, are expressed either in E. coli or in S. cerevisiae. IFN-β is expressed in E. coli and in CHO cells. In this latter case, the final products even differ structurally. While the CHO-produced IFN-β is amino acid-wise identical to the human original, the E. coli produced IFN-β is one N-terminal amino acid shorter; has a cysteine substituted by a serine, and lacks any glycosylation. So, interestingly, it is apparently not possible to anticipate superiority for the more “authentic” or “natural” version of the glycoprotein. A head-to-head clinical study, which compared the “artificial” IFN-β-1b and the “authentic” IFN-β-1a showed superiority in favor of the “artificial” IFN-β-1b [5]. However, definite proof, that the one IFN molecule is indeed superior to the other cannot be deduced from this study since both drugs were given in different doses and at different dosing intervals. IFN-β-1a induced less antibody formation and could therefore theoretically be regarded as better tolerated by the human body; however, improved safety has not been shown in clinical practice so far. Currently, there are three approved drugs produced using BHK cells, NovoSeven® (eptacog a), Helixate Nexgen® (Factor VIII), and Kogenate Bayer® (Factor VIII).

Additional production platforms ([6] and citations therein) for recombinant therapeutic proteins are now used or will be used in the near future (Table 1). Drugs produced in mouse myeloma cell lines (e.g., several recombinant antibodies) or even in the mammary gland of a living goat (anti-thrombin III), are already approved by the leading drug agencies (FDA, EMEA). In developing countries, the first recombinant preparations are being developed or even approved by national authorities, which are produced in plant systems (tobacco, moss, carrot), in insect cell system (Sf9), or in human HEK293 or PERC.6 cells, and others are likely to follow [7, 8]. None of these expression systems has general advantages in all situations and for all applications, and the choice for an expression system has to be made on a case-by-case basis, sometimes with rather unexpected outcomes. Epoetin delta, for example, which is produced “authentically” in a human fibrosarcoma cell line, seems to cause notably more cardiac events than the other approved epoetins, which are produced in CHO cells and are therefore slightly differently modified as compared to the human epoetin [9].

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 favorable, such as CHO or BHK cells. Lately, attempts have been made to modify other platforms like S. cerevisiae or plants by metabolic engineering [10, 11]. Such modified organisms are capable of introducing human-like sugar moieties.

**Authentic versus modified recombinant products**

Authenticity of recombinant drugs is no longer regarded as an important advantage. In fact, only a minority of all approved recombinant drugs correspond strictly to their “authentic” counterparts.

Besides the fact that early products, which were produced in E. coli, differed from the “original” due to technological concessions (e.g., lack of glycosylation, substitution of cysteine by serine, N-terminal amino acid addition or N-terminal amino acid loss), an important driver for intentionally introduced structural changes are the pharmacokinetic characteristics. Regardless of the choice of production platform, we are still far away from applying (and may probably never be able to apply) recombinant proteins in exactly the same way as the body delivers them. Insulin, for example, is subtly secreted from the pancreatic beta-cells upon physiological stimuli, whereas it is administered by a syringe subcutaneously in a very high millimolar concentration. This concentration forces recombinant insulin into hexameric aggregates that are unable to bind to the receptor. Receptor binding can only occur after dilution by diffusion, with the consequence that patients are advised to anticipate an efficacy lack for 15–30 min post injection. This problem, which clearly interferes with patient compliance, was elegantly solved by slight structural modifications at the C-terminal region of the B-chain of the protein core. The resulting second-generation fast-acting insulins were insulin lispro (Humalog®), insulin aspart (Novolog®) and insulin glulisin (Apidra®). These fast acting insulins are functionally complemented by long-acting insulin
Table 1. Characteristics of production platforms for recombinant therapeutic proteins\(^a\)

<table>
<thead>
<tr>
<th>Expression system</th>
<th>Classification</th>
<th>Development of system</th>
<th>Disulfide bonds</th>
<th>Glycosylation</th>
<th>Secretion</th>
<th>Cost of fermentation</th>
<th>Use of antibiotics</th>
<th>Safety costs</th>
<th>Processes developed</th>
<th>Product on market</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em></td>
<td>Gram-negative bacterium</td>
<td>Completely developed</td>
<td>(Yes) in the periplasm</td>
<td>No</td>
<td>Periplasmic secretion</td>
<td>Promoter-dependent low to moderate</td>
<td>Typically required</td>
<td>Low costs</td>
<td>Industrial scale</td>
<td>Yes</td>
</tr>
<tr>
<td><em>Saccharomyces cerevisiae</em></td>
<td>Budding yeast</td>
<td>Completely developed</td>
<td>Yes</td>
<td>Yes; high mannose</td>
<td>Possible</td>
<td>Low</td>
<td>Not required</td>
<td>Low costs</td>
<td>Industrial scale</td>
<td>Yes</td>
</tr>
<tr>
<td><em>Pichia pastoris</em></td>
<td>Methylo-trophic yeast</td>
<td>Completely developed</td>
<td>Yes</td>
<td>Yes; no terminal α1,3 mannose</td>
<td>Possible</td>
<td>Low</td>
<td>Not required</td>
<td>Low costs</td>
<td>Industrial scale</td>
<td>Yes</td>
</tr>
<tr>
<td><em>Hansenula polymorpha</em></td>
<td>Methylo-trophic yeast</td>
<td>Completely developed</td>
<td>Yes</td>
<td>Yes; no terminal α1,3 mannose</td>
<td>Possible</td>
<td>Low</td>
<td>Not required</td>
<td>Low costs</td>
<td>Industrial scale</td>
<td>Yes</td>
</tr>
<tr>
<td><em>Yarrowia lipolytica</em></td>
<td>Dimorphic yeast</td>
<td>Early stage</td>
<td>Yes</td>
<td>Yes; exact features jet unknown</td>
<td>Possible</td>
<td>Low</td>
<td>Not required</td>
<td>Low costs expected</td>
<td>Lab scale</td>
<td>No</td>
</tr>
<tr>
<td>Plant cells</td>
<td>Higher eukaryote</td>
<td>Completely developed</td>
<td>Yes</td>
<td>Yes; terminal fucose</td>
<td>Possible; size-restrictions</td>
<td>Moderate</td>
<td>Not required</td>
<td>Low costs</td>
<td>Pilot scale; production scale</td>
<td>Yes (Cuba)</td>
</tr>
<tr>
<td>Mammalian cells (e.g., CHO)</td>
<td>Higher eukaryote</td>
<td>Completely developed</td>
<td>Yes</td>
<td>Yes; typically human-like</td>
<td>Usually</td>
<td>High</td>
<td>Not required</td>
<td>High costs</td>
<td>Industrial scale</td>
<td>Yes</td>
</tr>
<tr>
<td>Animals</td>
<td>Mammals</td>
<td>Completely developed</td>
<td>Yes</td>
<td>Yes; typically human-like</td>
<td>Usually</td>
<td>Farming; moderate costs</td>
<td>Not required</td>
<td>High costs</td>
<td>Industrial scale</td>
<td>Yes</td>
</tr>
</tbody>
</table>

\(^a\) Modified from [6].
derivatives, again based on intelligent structural modifications. It is highly plausible that combinations of such modified insulins will prove beneficial in the long term, not as much based on molecular superiority, but based on superior patient compliance [12].

Another example of deliberate change from “authenticity” is tenecteplase (a modified tissue-specific plasminogen activator, TNKase®). In tenecteplase, the plasminogen activator inhibitor-1 binding site Lys-His-Arg-Arg was replaced by an Ala-Ala-Ala-Ala sequence. This resulted in a significantly longer biological half-life compared to native t-PA due to greater resistance to biological inactivation [13].

Similarly, reteplase (Rapilysin®) is a N-terminal truncated alteplase version; in addition, it is not glycosylated since it is produced in E. coli. This second generation alteplase was approved on the basis of favorable pharmacokinetic considerations allowing bolus therapy [14], although the authentically produced human alteplase (Actilyse®) was already on the market.

We will see a great variety of third generation biotechnology products as fragments from authentic human proteins in the future, many of them being antibody fragments or antibody-derived peptides.

Post expression modification of recombinant proteins

Not only modifications in the polypeptide chain but also in the post expression modification can provide the basis for second-generation biotechnology products. Modifying proteins with PEG chains, or with additional sugar chains, changes pharmacokinetic properties significantly, and frequently in favor of overall therapeutic efficacy. For example, PEGylated IFN-α variants, with a lengthened plasma half-life, clearly proved to be therapeutically superior in treating chronic hepatitis C and B infections compared to unmodified IFN-α [15].

Darbepoeitin (Aranesp®) with two additional glycosylation chains has a lower receptor affinity compared to unmodified erythropoietins but a significantly longer biological half-life. The overall therapeutic efficacy appears to be superior for the modified erythropoietins [16]. Darbepoeitin has a higher sialic acid content, providing more protection from hepatic uptake by the asialoglycoprotein receptors present on liver cells, as the result of the replacement of two amino acids by asparagine residues, thereby creating two additional sites for N-linked glycosylation. Darbepoeitin is produced in CHO cells, and its excellent safety record associated with repeated CHO platform to produce adequate administration of approved doses in patients [17] provides support to the value of the glycosylated human proteins, including those with high(er) sialic acid contents.

In the case of glucocerebrosidase, therapeutic efficacy turned out to be dependent on a complex sugar modification. Glucocerebrosidase is a lysosomal enzyme, which is usually delivered through an intracellular delivery system starting at the Golgi apparatus and ending in the lysosomes. Gaucher disease is characterized by an inborn deficiency of this enzyme and it needs to be infused intravenously to reach macrophages (the main site of pathological lysosomal accumulation of glucocerebroside) in sufficient quantities. Surprisingly, this could not be accomplished with the natural form of glucocerebrosidase but, after its sugar chains were cut back to expose mannose, such modified glucocerebrosidase was actually able to achieve therapeutic concentrations in target cells (monocytes and macrophages). Therefore, a complex biochemical modification of recombinant glucocerebrosidase sequentially using three specific glycosidases (neuraminidase, β-galactosidase and N-acetyl-glucosaminidase) is required to obtain mannose-terminated glucocerebrosidase (imiglucerase, Cerzyme®), which has proven to be a very safe and highly successful therapy in patients suffering from Gaucher disease [18]. These interesting observations, which elegantly solved a long-existing unmet medical need, cannot be extrapolated towards even higher efficacy by adding additional mannose groups. In a recent preclinical study, various expression and modification systems were used to produce different mannose-terminated glucocerebrosidases. The ones containing more mannose, and as such resembling proteins produced in yeast expression systems, were not more active than standard Cerzyme® [19].

Whereas mostly protein glycosylation will positively affect pharmacokinetic properties of a therapeutic protein, in the case of recombinant antibodies glycosylation may also negatively affect their activity [20]. The chimeric mAb rituximab, for instance, was differently active with respect to antibody-dependent cellular cytotoxicity based on the glycosylation state [21].

Immunological considerations

A major concern for all recombinant drugs is immunogenicity. It can be assumed that all biotechnology-produced therapeutics may exhibit
some form of immunogenicity during an unpredictable time scale. An antibody response can result in reduced efficacy of the therapeutic protein itself, or, very rarely, in inactivation of native endogenous proteins by neutralizing antibodies or in anaphylactic reactions [22, 23]. Well-known is the Eprex® case [24]. Between 1998 and 2002, unusual high numbers of pure red cell aplasia were reported in patients with chronic kidney disease who received subcutaneously administered Eprex, a particular recombinant human erythropoietin preparation (epoetin-α; OrthoBiotech, a division of Janssen.Cilag, marketed outside the US). Pure red cell aplasia resulted from depletion of endogenous and therapeutically applied erythropoietin by neutralizing antibodies induced by the treatment. Most likely, the presence of leachates in the formulation, arising from an interaction between polysorbate 80 and the uncoated rubber stopper were responsible for the observed antigenicity.

This highlights the fact that safety of new therapeutic protein products is difficult to predict. On the one hand, we have seen that combination of the recombinant protein and formulation/packaging constituents can be more critical for safety than changes in the recombinant protein itself. On the other hand, the huge number of molecular changes compared to the natural originals in authorized biopharmaceuticals and their clinically demonstrated fairly safe therapeutic use has put immunogenicity with respect to safety concerns in perspective.

Overall safety and in particular immunogenic safety, can only be assessed through clinical and post-marketing programs due to (i) the human immune system being more sensitive than the available physical tests or bioassays; (ii) the limitations of current analytical methods; (iii) the lack of standardized assays; and (iv) a certain degree of idiosyncrasy of each individual’s immune system and immune disposition. Rare immune-mediated reactions (e.g., 1 in 10 000 patient-years) will only become apparent through robust post-marketing surveillance.

Not yet adequately addressed is a potential reduction of efficacy of the therapeutic protein caused by neutralizing antibodies. If this is relevant, then this problem might be overcome by dosing consistently on the basis of antibody titers and not only on body weight [25]. Authorities are becoming more and more aware of this potential problem and may ask for studies in this direction.

Conclusions

The number of approved recombinant drugs now exceeds 120. Based on the unique potential of biotechnology to copy any genetic information in any living organism or cell, the first molecules were indeed exact (or almost exact) copies of human proteins produced in insufficient quantities in particular patients (substitution drugs). However, after scanning the list, it is obvious that the landscape has really changed. Recently developed biotechnologicals are exact copies of human proteins only in a very few cases. One of the biggest surprises is that these modified biotechnologicals are also fairly well tolerated and mostly do not provoke immune response complications. This is good news since modifications are frequently needed to improve pharmacokinetic properties of biotechnologicals.

Conflict of interest statement: Not relevant to the article, but in the interests of full disclosure, the author declares that he has received in the past honoraria for oral presentation from the following Biotech companies: Amgen, Genzyme, Merck-Serono, Roche, Stada and Wyeth.

References


