Therapeutic Glycoprotein in Mammalian

A Case Study of production

Dr. Jyoti Jyotsna
Lecturer in Biology
+2 Zila School, Darbhanga, Bihar - 846004

Abstract: The biopharmaceutical industry has significantly turned its biologics production towards mammalian cell expression systems. The presence of glycosylation machineries within these systems, and the fact that monoclonal antibodies represent today the vast majority of new therapeutic candidates, has largely influenced this new direction. Recombinant glycoproteins, including monoclonal antibodies, have shown different biological properties based on their glycan profiles. Thus, the industry has developed cell engineering strategies not only to improve cell’s specific productivity, but also to adapt their glycosylation profiles for increased therapeutic activity. One way to favor human-like glycosylation would be to use human cell lines for recombinant protein production. This strategy would warrant that proteins harbor, if not the ideal glycosylation pattern, at least a non-immunogenic glycans. The most commonly used human cell lines to manufacture glycoprotein therapeutics are the HEK293 cells and the HT-1080, respectively from human embryo kidney and fibrosarcoma origin.

Index Terms - Glycoprotein, Mammalian, Cells, Production, Therapeutic.

Introduction

While a major trend in the last decade has been to use mammalian cell lines to manufacture recombinant glycoproteins, it is not to be forgotten that there are still a large number of recombinant biotech products produced in other expression systems. How, these organisms do not have the ability to adequately glycosylate recombinant proteins, due to the absence of the required enzymatic machinery. These systems are thus mainly limited to the expression of non-glycosylated proteins. Bacterial expression systems have rapid cell growth and high yields, but proteins often aggregate and have to be extracted from inclusion bodies, due to the absence of chaperone proteins, before being refolded in vitro. Still, some commercially available enzymes that are not glycosylated, such as asparaginase and collagenase, are produced in bacterial expression systems. Some recombinant proteins are produced in yeast, which can also divide rapidly and generate high yields. However, these cells produce glycoproteins with high-mannose glycans structures, which may be immunogenic and less potent in humans. Examples of approved therapeutics from yeast expression system are ceriplasmin (JETREA) and caridecastog (TRETEN). As for plant and insect cells, both are able to produce recombinant proteins with complex glycans, but with structures quite different from the human ones. Indeed, plants produce core α1,3-fructose and β1,2-xylose, which are completely absent in human cells and could be immunogenic. Insect cells produce N-glycan precursors that are trimmed, creating high mannose or paucimannose structures. Both plant and insect cells lack sialic acid residue on their glycans. Several attempts have been made to glycoengineer plant and insect cells for protein productions. In 2012, a first plant-generated therapeutic recombinant protein, taliglucerase alfa was approved by the FDA. Currently, the only therapeutics approved for insect cell expression systems are the human papillomavirus vaccine, the prostate cancer immunotherapy vaccine and the flu vaccine. Lastly, a few therapeutic proteins are produced in transgenic animals. As with other mammalian expression systems, transgenic animals often produce different glycosylation patterns compared to native human proteins. The first therapeutic produced in transgenic animal available on the market was the human anti-thrombin alpha, made in transgenic goat milk. This product was followed by the C1-esterase inhibitor produced in rabbit milk, approved by the EMA in 2011 and by the FDA in 2014. Finally, a third product generated in transgenic chicken eggs has been approved in 2015 by both EMA and FDA, a recombinant human lysosomal acid lipase.

More than a hundred new biopharmaceutical products have been approved and marketed in the Asia and European Union. Market value for these biologics was recently estimated at $140 billion US, with a total of over two hundred therapeutics. A significant portion of these products are recombinant proteins, with an ongoing increase in the number of them produced in mammalian expression platforms. This trend is mostly driven by the increased attention directed to post-translational modifications of these biologics, in particular towards their glycosylation state. Indeed, several efforts have been made over the last few years to understand how glycosylation can influence the biological activity of therapeutics. Studies have demonstrated that proper glycosylation profiles can improve recombinant protein properties such as increase their stability and half-life in blood circulation and decrease their immunogenicity. Among the mammalian-based expression systems, CHO cells is by far, the most commonly used cell line. It is involved in the production of over 70% of recombinant biopharmaceutical proteins, most of them being monoclonal antibodies. This review will summarize the
recent advances in production of glycoproteins in mammalian cells, with a particular emphasis on the CHO cell system. The various expression systems currently used for therapeutic glycoprotein production will be overviewed and cell engineering strategies used to improve biologics production and/or quality will be discussed. Finally, we will also describe the different “omics” approaches used lately in the field in order to improve glycoprotein production and/or glycosylation.

OVARY CELLS

The human immune system can produce antibodies against these N-glycans that could contribute to immunogenicity/neutralization of the corre-sponding biotherapeutics. CHO cells also have limited ability to gamma-carboxylate recombinant proteins such as clotting factors, even though some improvements have been achieved through metabolic engineering work. Proteins requiring proteolytic processing for maturation may not always be fully cleaved and active when expressed in CHO. For example, co-expression of furin was shown to allow the production of fully cleaved and active von Willebrand factor in an industrial-scale CHO perfusion system and of the coagulation factor VIII B-domain. Similarly, co-expression of proprotein convertases allowed for the efficient maturation of human bone morphogenetic protein-7. CHO cells are widely used for glycoprotein production because of their numerous advantages. These cells can achieve substantial produc-tion rate, are suitable for large-scale industrial suspension culture and can be adapted to grow in various serum-free and chemically defined culture media. Since CHO cells produce recombinant glycoproteins with human-like glycans, the generated products are to more likely be compatible and bioactive within human hosts. Furthermore, these cells are refractory to infection by human viruses, which mini-mizes biosafety risks for commercial production purpose.

This decreased susceptibility could be attributed to the fact that many viral entry genes are not expressed in CHO cells. Moreover, different gene amplification systems have been developed and used in CHO cells, which allow for high titer yields and good specific productivity. There are many examples of biotherapeutic glycoproteins approved by the Food and Drug Agency and the European Medicines Agency currently produced in these cells, Several monoclonal antibodies such as Siltuximab, Pertuzumab and Rituximab, as well as other proteins such as tissue plasminogen activator and Human DNase are just some of the many examples of biotherapeutics generated in CHO cells. In 2015, more than half of the thirteen new biologics approved were recombinant proteins produced in CHO cells. Among these products, four monoclonal antibodies, Duratahumab, Mepolizumab and Evelocu-mab/Alirocumab are used to treat multiple myeloma, asthma and hypercholesterolemia, respectively. The same trend is currently observed in 2016, where again more than half of the approved biotherapeutics are produced in CHO cells. Although CHO cells possess many advantages for glycoprotein produc-tions, they are unable to produce some types of human glycosylation, such as α-2,6-sialylation and α-1,3/4-fucosylation. Moreover, CHO cells produce glycans that do not occur in human cells, namely N-glycolylneuraminic acid (Neu5Gc) and galactose-α1,3-galactose (α-gal), even though these occurring at very low levels (e.g < 2% and < 0.2% respectively).

CELL LINES OF HUMAN

Approvals for therapeutic proteins produced in human cell lines was observed, with four new FDA/EMA authorized glycoproteins, rFVIIIFc and rFIXFc are two of these proteins, used for prevention of bleeding episodes in people with hemophilia A and B. They consist of domains of FVIII and FIX proteins fused to the Fc portion of immunoglobulin G1 (IgG1). rFVIIIFc has six tyrosine sulfation sites which are essential for its functionality. Besides, rFIXFc has γ-carboxyl- ylation of its first twelve glutamic acid residues, also important for its activity. Expressing these glycoproteins in HEK293 resulted in greater tyrosine sulfation and glutamic acid γ-carboxylation compared to CHO cells and excluded any α-gal and Neu5Gc from the manufactured products (Berkner, 1993; Kannicht et al., 2013; Peters et al., 2013; McCue et al., 2014, 2015). Dulaglutide, another Fc fusion protein used for treatment of type 2 diabetes mellitus, is produced in HEK293-EBNA1 cells and was approved in 2014. Lastly, the Human-cl rhFVIII, a replacement clotting factor for hemophilia A disorder is approved by EMA since 2014 and by FDA since 2015. It is produced in HEK293-F cell line and has shown similar glycosylation profile to the plasma-derived factor VIII, deprived of α-gal and Neu5Gc. One way to favor human-like glycosylation would be to use human cell lines for recombinant protein production. This strategy would warrant that proteins harbor, if not the ideal glycosylation pattern, at least a non-immunogenic glycans.

The most commonly used human cell lines to manufacture glycoprotein ther-apeutics are the HEK293 cells and the HT-1080, respectively from human embryo kidney and fibrosarcoma origin. Drotrecogin alfa, the first therpeutic glycoprotein produced in human cells (HEK293) approved by FDA and EMA, was accepted by both agencies in 2001 and 2002 respectively. However, it was removed from the market in 2011, since it failed to show significant beneficial effects. Yet, only four biological glycopro-teins were approved in the following decade by FDA and/or EMA. These four therapeutics, namely Agalsidase alfa, Epoetin delta, Idursulfase and Velaglucerase alfa, are produced using a gene activation technology (proprietary of Shire) in HT-1080 cells (Moran, 2010). Epoetin delta produced in HT-1080 was found to have better homogeneity of its tetra-antennary glycans, higher sialic acid content and no Neu5Gc, compared to CHO-produced erythropoietin. However, this product was volun-tarily withdrawn for commercial reasons. As for Velaglucerase alfa, its glycoprofile has also been compared to similar products (other β-glucocerebrosidases) produced in CHO cells and carrot cells. Even though these three products show diverse glycans profiles, they showed similar macrophage uptake, in vitro enzymatic activity, stability and efficacy. Notably, neutralizing anti-IgG1 or IgG2 antibodies were observed in 24% of patients, with an impact on the protein activity.

Some human cell lines are currently being used in preclinical and/or clinical development stages for recombinant glycoprotein production. This is the case for PER.C6 cells, which consist of human embryonic retinoblasts transformed with adenovirus type 5 E1A and E1B-encoding sequences. These cells are able to produce high titers of IgG without requiring amplification of the incorporated gene. MOR103 is a mAb directed against granulocyte macrophage colony-stimulating factor and is developed to treat patients with rheumatoid arthritis and multiple sclerosis. CL184 is a combination of two mAbs used against rabies virus (Marissen et al., 2005; Bakker et al., 2008). Both antibodies, produced in PER.C6 cells, are currently tested in clinical phase 1/2 (Dumont et al., 2015). The HKB-11 cell line is a
fusion of HEK293S and human B-cell lines. It has recently showed high-level protein production and α2,3 and α2,6-sialic acid linkages. Two other cell lines, the CAP cells of human amnionocytes origin and the HuH-7 cells of human hepatocellular carcinoma origin, are presently tested for recombinant glycoprotein production in preclinical phases and both display human-like glycosylation profiles.

**CELL LINES OF MAMMALIAN (NON HUMAN)**

Murine myeloma cells (NS0 and Sp2/0), derived from tumor cells that no longer produce their original immunoglobulins, also are being used to produce some commercial monoclonal antibodies such as Cetuximab and Palivizumab. Baby Hamster Kidney (BHK) cells are mostly been used for the production of vaccine. Only two mar-keted recombinant glycoproteins are currently manufactured in these cells, Factor VIII and Factor VIII, which are other clotting factors. These large glycoproteins are abundantly glycosylated and sulfated and are thus challenging to manufacture. In 2015, three new therapeutic monoclonal antibodies produced in murine cells were approved by the FDA, namely Dinutuximab, Necitumumab and Elotuzumab, all used to treat different cancer types. Murine cells can also produce α-gal and Neu5Gc at considerably higher levels than hamster cells, increasing the risks of immunogenicity.

**EXPRESSION SYSTEMS**

While a major trend in the last decade has been to use mammalian cell lines to manufacture recombinant glycoproteins, it is not to be forgotten that there are still a large number of recombinant biotherapeutics produced in other expression systems. However, these organ-isms do not have the ability to adequately glycosylate recombinant proteins, due to the absence of the required enzymatic machinery. These systems are thus mainly limited to the expression of non-glycosylated proteins. Bacterial expression systems have rapid cell growth and high yields, but proteins often aggregate and have to be extracted from inclusion bodies, due to the absence of chaperone proteins, before being refolded in vitro. Still, some commercially available enzymes that are not glycosylated, such as asparaginase and collagenase, are produced in bacterial expression systems. Some recombinant proteins are produced in yeast, which can also divide rapidly and generate high yields. However, these cells produce glycoproteins with high-mannose glycans structures, which may be immunogenic and less potent in humans. Examples of approved therapeutics from yeast expression system are ocrilplasmin and catridecagog. As for plant and insect cells, both are able to produce recombinant proteins with complex glycans, but with structures quite different from the human ones. Indeed, plants produce core α1,3-fructose and β1,2-xylene, which are completely absent in human cells and could be immunogenic.

Insect cells produce N-glycan precursors that are trimmed, creating high mannose or paucimannose structures. Both plant and insect cells lack sialic acid residue on their glycans. Several attempts have been made to glycoengineer plant and insect cells for protein productions. In 2012, a first plant-generated therapeutic recombinant protein, taliglucerase alfa was approved by the FDA. Currently, the only therapeutics approved for insect cell expression systems are the human papillomavirus vaccine, the prostate cancer immunotherapy vaccine and the flu vaccine. A few therapeutic proteins are produced in transgenic animals. As with other mammalian expression systems, transgenic animals often produce different glycosylation pattern compared to native human proteins. The first therapeutic produced in transgenic animal available on the market was the human anti-thrombin alpha, made in transgenic goat milk.

In the biopharmaceutical industry, production of glycoproteins is currently achieved by either transient or stable gene expression in the cells. When the need for a quick and economical approach prevails, transient expression remains the best choice for protein production. By skipping the lengthy selection process for the cells that have integrated the plasmid within their genome, transient transfection is much faster. However, the production rate relies on many factors including efficiency of the transfection phase, cytotoxicity of the transfection reagent and exten- siveness of the feeding strategy applied to the culture post-transfection. Although the level of protein obtained is not as high as with stable gene expression, it is still sufficient for many applications. Indeed, during transient transfection, the plasmid DNA is mostly kept extrachromosomally, the cells rapidly losing it during division, therefore limiting the amount of expressed proteins. Yet, it is ideal for high throughput screening for hits identification and very helpful for early stage product characterization. To date, only viral vectors used in gene therapy have been produced by transient transfections for clinical applications. When it comes to production of glycoproteins in large quantities, stable gene expression systems remain the preferred avenue. For these systems, multiple aspects have been tackled and optimized for improving productivity, process robustness and reducing cell line generation timelines.

**PROCESSING OF SELECTION SYSTEMS**

NS0 and Sp2/0 cell lines express insufficient endogenous levels of GS to support growth, therefore simply removing glutamine from the culture media is sufficient for selection. However, for CHO cells, it is necessary to add methionine sulfoximine MSX in combination with the removal of glutamine in the media, to help inhibit endogenous glutamine synthetase activity and have enough selection pressure. Moreover, CHO GS-Knock out (KO) cell lines were recently developed (Eli Lilly), thus enhancing the stringency of this selection system. Many selection systems have been developed over the years to improve the production rate of the proteins and the efficiency of the selection. For stable expression, a gene marker is usually integrated in the expression plasmid along with the cDNA encoding the gene of interest, conferring a selective advantage to the cells that integrate the plasmid into their genome. Copy number of integrated plasmid as well as the integration site(s) within the host genome are some of the key factors when it comes to stable gene expression. There are two frequently used selection markers in the biopharmaceutical industry, namely the glutamine synthetase (GS) and the dihydrofolate reductase genes.

The first one, originally developed by Celltech (now Lonza Biologics), uses complementation of a glutamine auxotrophy by a recombinant GS gene. Similarly, the DHFR selection uses the fact that CHO cell mutants have been selected to be deficient for DHFR...
(Urlaub and Chasin, 1980). Using a recombinant DHFR gene in the integrated plasmid, cells are put under selection pressure and gene amplification with increasing concentrations of methotrexate to inhibit the DHFR enzyme activity, combined with the absence of nucleotide precursors in the media. Other selection systems have been developed more recently, such as the OSCAR™ system from the University of Edinburgh, that uses minigene vectors encoding the hypoxanthine phosphoribosyltransferase (HPRT), essential for purine synthesis. Yet, only the DHFR and the GS systems have been used so far by the industry for high-scale commercial productions.

Despite the success of the selection systems described above, a main problem persists with these methods, being the fact that they are based on random plasmid integration and its expression cassette in the host genome. This random integration creates very heterogeneous cell population that have variable expression levels amongst the different clones. Transgenes are likely to be inserted in heterochromatin regions, which results in very weak gene expression levels. This implies screen- ing a very large number of clones (generally multiple hundreds to a few thousands, depending on the cloning method used) in order to find those rare ones with stably integrated plasmid in highly transcription- ally active chromatin regions (“hotspots”). Recently, several molecular and cellular biology tools have been developed for targeted site gene integration. This approach could eventually help biopharmaceutical companies minimize the randomness of gene insertion and increase the predictability for high transgene expression. Although tremendous efforts are required to isolate a highly productive and stable recombination host cell line compared to random integration strategies, it remains a very attractive avenue that should help control and predict the expression level of the daughter clones. Since the transgene and selection marker are on the same plasmid, their co-integration within the same chromatin region increases the odds that good expression of the marker will also result in good transgene expression.

As first generation tools, several recombinase systems have been used for targeted site approaches, among which the Cre/Lox and the Flp/FRT were mostly utilized. Recombination mediated cassette exchange (RMCE) technology is attracting interest by the industry for targeted gene insertion. It uses a site specific recombinase to exchange a cassette flanked by heterospecific sequences in a plasmid with a cassette flanked by identical sequences within the host genome. This technology had good success in increasing the success rate and reducing timelines for the generation of stable industrial-grade CHO cell lines stably expressing monoclonal antibodies.

As second generation tools, different endonucleases have been employed. This category includes zinc finger nucleases (ZFN), meganucleases, transcription activators like effectors nucleases (TALEN) and CRISPR/Cas9. These nucleases induce DNA double-strand breaks at precise locations into the host genome, facilitating the integration of the cassette by non-homologous end-joining (NHEJ) or homology-directed repair. Both ZFN and TALEN technologies rely on the ability to customize a DNA-binding domain for a specific sequence (the targeted sequence for cleavage) combined to a nuclease effector domain. These nucleases are thus dependent on the ability of developing a good DNA recognition motif, with high specificity and affinity. Because the genome comprises numerous sequence repetitions or has highly homologous DNA sequences, efforts have been made in the last years to improve the off-target effects and the specificity of these enzymes. As an efficient alternative to these protein-based genome editing nucleases, the CRISPR/Cas9 technology has just emerged. It relies on a RNA-guided cleavage with a CRISPR RNA that will bind to a specific DNA sequence (seed) combined with the Cas9 endonuclease. This technique has already been validated in CHO cells and reduces the production variability between clones. ZFN, TALEN and CRISPR/Cas9 technologies are however mostly used for gene-specific knockout.

Some cis-acting epigenetic regulatory elements have also been developed to improve the production level and stability of the producing cell lines. These elements help remodelling the chromatin environment to maintain an active transcriptional state around the transgene. One of the most frequently used cis-acting elements is the scaffold/matrix attachment region (S/MAR). Many reports have shown that including these MAR elements in expression vectors for recombinant proteins could significantly improve the expression levels. However, conflicting results in the literature have led to the conclusion that MAR activity could be sequence-specific and is influenced by specific vector configuration. In recent years, development of bioinformatic tools capable to predict the outcome of MAR sequences helped to improve MAR elements.

The ZFN knockout approach was particularly successful for specific deletion of the GS and the DHFR genes in CHO cells, thus improving the selection stringency of the generated cell lines. Further applications related to gene knockout for glycoprotein production will be discussed in the genetic engineering section. Even if these tools could prove very useful for site-specific integration, one of the key challenges is still to identify good hotspots in the host genome that will allow good expression levels and stability. One should also keep in mind that such specific integration may not be a one-size-fits-all approach as some therapeutic proteins may require a particular level of expression to fold properly, or to acquire adequate quality attributes (e.g. glycosylation, proteolytic processing, etc.) Another, but less commonly used tool, is the mammalian artificial chromosome expression (ACE) technology. This minigenome serves as an autonomous genetic element that replicates with the cells.

Its DNA sequence is customizable with various regulating elements that could possibly help for its expression. It was shown to be effective for establishing CHO cell lines producing respectable titers of monoclonal antibody. Fed-batch performance and stability of ACE vs random integration systems have been compared for IgG1 expression in CHO cells and have shown similar performance. Also, the PiggyBac™ (System Bioscience Inc.) transposon system uses an efficient transposase purified from the cabbage looper (Trichoplusia ni) to easily integrate the gene of interest into the host genome (Ding et al., 2005). This approach has recently shown improved yields for stable production of antibodies in CHO cell lines (Rajendra et al., 2016).

Another class of epigenetic regulatory elements is the ubiquitously active chromatin opening elements (UCOEs). These sequence contain CpG rich islands found within the promoter regions of housekeeping genes which confer an open chromatin state for transgene expression. UCOEs were also found to be useful for increasing the productivity of recombinant protein cell lines. The third class of epigenetic elements is the stabilizing and anti-repressor (STAR) element. They were discovered by a genomic screen of elements that would increase transgene expression.

Inclusion of these various epigenetic elements not only can help to increase the expression level of biotherapeutics, but can also increase the number of clones that have integrated the transgene with a more defined copy number of transgene per cell, thus accelerating the selection process. It was also demonstrated that MAR and UCOE helped decrease the variability of expression between the different
clones and contributed in maintaining an active transcriptional state for better transgene expression. It has been shown that an UCOE can also prevent DNA methylation of the CMV promoter region, thus preventing gene expression silencing. Since most of these elements are composed of nucleosome depleted regions, this could also contribute to the inhibition of the propagation of an inactive chromatin state surrounding the transgene, serving as genetic boundaries.

**GROWTH AND SURVIVAL OF CELLS**

Producing more proteins with fewer resources and in less time remains a significant challenge when working with mammalian cell expression systems. In this perspective, various tools and methods focusing on cell growth, proliferation and survival have been tackled by the biopharma industry and academia. Key aspects for the optimization of cell growth and productivity certainly are the process parameters, as well as the media and the feeds composition. These media and feeds should maintain availability of key nutrients and avoid accumulation of metabolites having negative impact on the culture performance. However, the composition of most of the commercial media and feeds are industrial trade secrets, rendering their optimization resource-intensive. Beside media and feeds, numerous small molecules have been identified from a chemical library that could enhance protein production. Sodium butyrate (NaB) and valproic acid (VPA) are often used in cell culture media. These histone deacetylase inhibitors, by preserving histones acetylation, reduce their interaction with DNA and helps maintain chromatin in an open configuration, leading to enhanced or prolonged transgene transcription. Addition of these chemicals has resulted in increased productivity in numerous studies, especially in CHO cells. However, the use of these inhibitors may also have undesired secondary effects, such as blocking cell cycle or inducing apoptosis. To attenuate these effects, optimization of the addition timing and concentration has shown to be important. Moreover, combining these inhibitors with mild hypothermia (30–32 °C) during production phase or with anti-apoptosis cell engineering could also increase productivity.

These hypothermia and anti-apoptosis strategies have been used to extend cell survival of CHO cells and thus maximize productivity. Indeed, simply using reduced temperature during protein production phase in CHO cells significantly improved the yield of numerous model proteins such as alkaline phosphatase, erythropoietin, interferon-γ, and β-interferon by slowing down cell cycle progression and allowing cells to shift from a proliferative to a productive mode. This shift toward protein production favors sustained nutrient availability and reduced secondary metabolite accumulation. As for anti-apoptosis engineering, genes of the Bcl-2 family (Bcl-2, Bcl-xL and Mcl-1) were overexpressed in a variety of CHO cell lines. Oppositely, down-regulation of pro-apoptotic genes like Bax, Bak, microRNA mir-mmu-miR-466h-5p and caspase-3, -7, -8 and -9 were also tested. Some engineering strategies rather rely on cell proliferation control. For example, the E2F-1 cell cycle transcription factor overexpression contributes to increase viable cell density in CHO batch cultures. The CDKL3 gene insertion also accelerates cell proliferation of 293 and CHO cells (Jaluria et al., 2007). Cell cycle progression was also faster when using c-myec oncogene in CHO cells (Kuystermans and Al-Rubeai, 2009). The mTOR pathway, involved in cell proliferation, survival and translation has also been exploited. The introduction of the mTOR gene in CHO cells resulted in increased cell proliferation and productivity of therapeutic IgG, alkaline phosphatase and α-amylase. Moreover, the chemical manipulation of this pathway can also enhance glycoprotein production in these cells. Furthermore, overexpression of the mir-17 microRNA in CHO was shown to augment cell proliferation speed, leading to higher Epo-Fc yield.

However, playing with the level of ER proteins involved in the secretory pathway has shown mixed results, depending of the expression system and/or the overexpressed protein. Some reports have revealed positive effects in overexpressing XBP-1 (X-box binding protein 1), a transcription factor regulating the unfolded protein response (UPR) during ER stress (Iwakoshi et al., 2003; Tigges and Fussenegger, 2006; Becker et al., 2008; Ku et al., 2008). However there were no obvious benefit observed in stable CHO-K1 cells expressing IFNy, EPO or a human mAb (Ku et al., 2008). Also, in response to ER stress, cells normally attenuate their translational machinery. ATF4, another critical transcription factor in the UPR system, has been overexpressed to restore translation and enhance levels of anti-thrombin III produced in CHO cells. Another limitation of glycoprotein production is the proper folding of the desired biologics. As these proteins are being translated, they travel through the endoplasmic reticulum (ER) and the Golgi for proper glycosylation and folding before being secreted. It has been suggested that translational and/or post-translational processes might also be rate-limiting for protein production, because of the lack of correlation between gene copy number and the amount of secreted proteins.

This activation is made through the GADD34 protein, which recruits the PP1 phosphatase responsible for elf-2A translation initiation factor dephosphorylation and activation. Similarly, GADD34 can also be overexpressed in order to increase productivity of CHO cells. Additionally, secretory bottlenecks of CHO cells may be relieved by overexpressing soluble N-ethylmaleimide-sensitive factor attachment protein receptors (SNAREs). For example, by controlling vesicle docking and fusion, expression of SNAP-23 and VAMP8 has been shown to improve production of SEAP and monoclonal antibodies in CHO-K1 cells.

Similarly, overexpression of ceramide transfer protein (CERT) S132A mutant, which is responsible for the transfer of ceramide from the ER to the Golgi and for sphingomyelin synthesis, also improved production of t-PA, human serum albumin and monoclonal antibodies in industrial cell lines. Likewise, signal recognition particle 14 (SRP14), which targets secreted proteins to the endoplasmic reticulum membrane, enhanced productivity of recombinant IgG in CHO cells when overexpressed.

Chaperones play a key role in the folding process of nascent proteins. They have been employed by many groups as cell engineering tools but have given variable results. The most popular chaperone for CHO cell engineering is the protein disulfide isomerase (PDI), which catalyzes the formation of disulfide bonds in proteins. Its overexpression has shown positive results on productivity in the case of monoclonal antibodies in both CHO cells and NS0 cells. However, other studies observed either a negative or no effect on productivity (Mohan et al., 2008). Similarly, the heavy chain-binding protein (BiP) chaperone gave mixed results, with enhanced or decreased productivity. Moreover, overexpression of the endoplasmic-reticulum ERP57 protein, an isoform of PDI, together with calnexin/calreticulin was able to increase thrombopoietin production in CHO cells. Overexpression of the Hsp27 and Hsp70 chaperones in CHO cells also increased recombinant IFNγ production. To explain such discrepancies between studies, it was recently determined that CHOK1 hosts have larger endoplasmic reticulum and higher mitochondrial mass compared to DUXB11-derived CHO cells.
Thus, these various and sometimes contradictory results could be partly explained by the use of different CHO host strains. Other cellular pathways have been engineered to sustain cell productivity. Among these approaches, anti-autophagy engineering and metabolic engineering have been used in order to optimize recombinant protein production in mammalian cells (for a detailed review on the topic). Yet, most of these strategies still need to prove their effectiveness and robustness for increasing recombinant protein production in mammalian cells, especially at an industrial level.

As mentioned previously, the expression systems used for glycoprotein production have significantly different glycosylation machineries. Plants, yeasts and non-human cell lines can generate glycans that are often absent from endogenous human proteins. For example, yeasts and insect cells can produce high-mannose or pauci-mannose oligosaccharides, while plant cells introduce undesirable α1,3-fucose and β1,2-xyllose. Moreover, as opposed to human cells, CHO and BHK cells do not express any α2,6-sialyltransferase activity and can generate glycans with terminal Neu5Gc and α-gal. However, both Neu5Gc and α-gal levels are significantly higher in murine cells such as NS0 and Sp2/0, compared to hamster cells. To avoid any potential immunogenic response, glycans present on biopharmaceutical proteins should be compatible for human hosts. Over the years, many studies have helped elucidate the biological functions of protein glycosylation. From these studies, specific glycosylation profiles have been identified, which could increase stability and efficacy of therapeutic agents. Consequently, several glycoengineering strategies have emerged to recreate these beneficial profiles on recombinant proteins. In particular, afucosylation and sialylation strategies are now popular glycoengineering approaches for biologics.

Since the sialic acid pathway includes many steps, various glycoengineering strategies have been elaborated in order to improve sialylation of biologics. Some of these approaches are based on the supplementation of cell culture media with different precursors of the sialic acid pathway, such as CMP-sialic acid, ManNAc and NeuNAc, but their effects are limited and still debated. Transient expression of different sialylation machinery enzymes was also tested. A functional analysis of 31 glycosyltransferases was even conducted for human EPO glycosylation optimization in six different mammalian cell lines, where enhanced expression of ST3GaIII, ST3GalIV and ST6Gal1 could enhance sialylation in HEK293, Cos-7, 3T3 and NSO cells. Transient expression of ST6Gal1 in CHO and HEK293 cells also increased the α2,6 sialylation of a trastuzumab F243A mutant antibody. Co-expression of ST6Gal1 and β4GalT1 in CHO cells also significantly enhanced the sialylation level of the mutated version of trastuzumab (Raymond et al., 2015). Co-expression of the same enzymes in HEK293 also helped optimizing sialylation of human Fc-FvIg (intravenous immunoglobulin), therefore enhancing its therapeutic activity (Washburn et al., 2015). Finally, transient transfection of the Chinese hamster ST6Gal1 gene in CHO cells expressing a bispecific antibody, also considerably increased the level of α2,6 sialylation.

As genome-editing strategies, the CHO cell lines producing recombinant lectin staining, a lectin which mainly contain. An engineered CHO cell line R26SL-R266Q glycosylation enzymes stably expressing the human α2,3-sialyltransferase ST3GaⅢ, the rat mutated GNE/MNK1 and the Chinese hamster CMP-sialic acid transporter was successful in increasing tetra-sialylation of recombinant human EPO. As mentioned earlier, the silkworm hemolymph anti-apoptotic protein 30K or 19 was shown to increase protein production. A CHO cell line stably expressing both 30Kc19 and human EPO showed that it could also significantly increase the EPO sialylation. In an effort to characterize the contributions of the various glycosylation enzymes in CHO-K1 cells, knockouts of 19 glycosyltransferase genes were recently achieved and EPO glycosylation was characterized to measure their impact. This study revealed predominant roles for some enzymes into CHO cells glycosylation pathway. Noteworthy, knockout of both ST3Gal4 and ST3Gal6 genes affected sialylation. Additionally, the same study presented a human ST6Gal1 knock-in experiment in ST3Gal4/6 knockout CHO cells, which resulted in considerable increase of α2,6 sialylation. Furthermore, to improve N-glycan homogeneity on therapeutics, a technique called “GlycoDelete” engineering was also recently developed in HEK293 cells. This system consists of engineered 293S GnTIII(−) cells stably expressing the catalytic domain of the endo-β-N-acetylglucosaminidase from the fungus HypoREA JECORINA fused to the Golgi targeting domain of human ST6Gal1. Using this expression system for transient anti-CD-20 production, recombinant proteins had very short sialylated N-glycans with high homogeneity on the Fc domain and reduced affinity to FcR compared to proteins expressed in wild-type 293S. This reduced affinity might be required to increase safety, for example in the context of neutralizing antibodies used for inflammatory cytokine targeting. Finally, there are also transgenic strategies to improve sialylation for other expression systems. The tobacco plant Nictiana Benthamiana expression system was genetically modified to carry the human-type sialylation pathway and to generate mono- and disialylated structures with α2,3 and α2,6 linkage (Kallolimath et al., 2016).

It was reported that afucosylated Fc domain glycans of human IgG1 could enhance the binding to FcγRIIa subtypes receptors of natural killer cells. Recently, the same effect was observed for human IgG4. This increased affinity allows better therapeutic activity of Mabs used in cancer therapy, by triggering the antibody-dependent cell-mediated cytotoxicity (ADCC) pathway. Since mAbs represent by far the most important class of therapeutic glycoproteins being manufactured, research towards development of strategies to produce low or non-fucosylated Mab glycans has intensified. As early strategies, CHO Lec13 cells and rat hybridoma YB2/0 cells were employed, the former for its impaired synthesis of GDP-fucose and the latter for its reduced expression level of the fucosyltransferase FUT8. Since then, some FUT8 knockout CHO cell lines have been established and one is currently available on the market for glycoprotein productions. The first FUT8 cell line was created by sequential homologous recombination. Recently, both ZFN and a CRISPR/Cas9 approaches were also used for FUT8 gene-specific knock-out. Disruption of the fucosyltransferase in these cell lines completely abolished the fucosylation on the Fc domain of IgG. Other methods to reduce the fucosylation level of antibodies glycan were also examined. Since expression of the β1,4-mannosyl-glycoprotein 4-β-N-acetylglucosaminyltransferase III (GnTIII) blocks fucosylation, CHO cells were engineered with both GnTIII and Golgi mannosidase II (ManII) for IgG production. Others have tried to reduce the fucosyltransferase level using siRNA (Imai-Nishiya et al., 2007). Heterologous expression of the bacterial enzyme GDP-4-dehydro-6-deoxy-β-mannose reductase (RMD) in CHO cells strongly reduced their fucosylation capacity. Lastly, other expression systems have been used to produce non-fucosylated antibodies. For example, a Pichia pastoris yeast strain was engineered to generate human-like glycans but devoid of core-fucosyl (Choi et al., 2003; Hamilton et al., 2006; Hamilton and Gerngross 2007). Consequently, the anti-CS1 antibody produced with
these cells are afucosylated and hasenhanced in vitro ADCC activity and in vivo anti-tumor activity compared to the same antibody produced in HEK293 cells. Addition of terminal sialic acids on the glycans of protein therapeutics helps to maintain them into the blood circulation by prevent- ing their recognition by the asialoglycoprotein receptors highly ex- pressed in liver hepatocytes. Thus, increasing the sialylation of proteins may decrease the frequency of injection or the amount of therapeutic protein used in a single dose. Therefore, it has become attractive for the industry to produce proteins with optimal sialylation as it could generate substantial savings and confer a therapeutical advantage for the patients.

In order to increase glycoprotein sialylation levels, strategies to inhibit sialidases have also been employed. The mammalian genomes possess 4 sialidase genes or pseudogenes (depending on the specie) which are expressed in various cellular compartments. Many chemical inhibitors have been synthesized to target these enzymes but their use in large-scale bioprocess productions is still very limited due to their elevated fabrication cost. Some attempts of knocking down the different CHO sialidases by shRNA or siRNA also improved the sialylation level of recombinant hIFNγ, particularly for Neu2 and Neu3 sialidases.

In addition to genomics, transcriptomics and proteomics tools, metabolomics approaches are now used in order to optimize glycoprotein production in CHO cell lines. Recent advances in metabolite quantification have allowed identification of cellular phenotypes under specific experimental conditions. Nutrient utilization and metabolic by-products accumulation are now easily quantifiable and serve as read-outs to improve cell culture conditions. Such tools have particularly contributed to the optimization of feeding cocktails and culture media to increase recombinant glycoprotein production and extend cellular growth. Moreover, new targets for cell engineering approaches can be identified, based on metabolomics profiling. A bottleneck at the malate dehydrogenase II (MDHII) level was characterized for the tricarboxylic acid (TCA) cycle in CHO cells and pyruvate metabolism was shown to vary between high producing and low producing anti-CD20 CHO clones. Finally, a multi-omics study combining transcriptomics and metabolo-mics data, identified variations in gene expression and in enzymatic reactions during the transition from a parental HEK293 cell line to a producer cell line.

**CONCLUSIONS**

These powerful tools will bring valuable contribution to the advance of research and development in biotechnology and lead to the next generation of cell factories. Mammalian cell lines, in particular CHO cells, are now extensively used for production of therapeutic glycoproteins by the biopharmaceutical industry. These cells possess many advantages in terms of cell culture and have the capacity to generate high titers.

Furthermore, the “omics” revolution brings up new challenges, as for scientists to integrate and analyze a tremendous amount of data. Considerable efforts in data mining and in the development of modeling tools will certainly be required. Yet, combining all “omics” data using computational models will help broaden our understanding and improve the various expression systems used by the biopharmaceutical industry for glycoprotein production. Over the years, many cell engineering strategies were attempted in order to increase such titers by optimizing selection markers, gene expression, cell growth and proliferation or protein folding and secretion. Among those engineered tools, CRISPR/Cas9 and RMCE technologies will largely contribute to the advance of glycoprotein production in a near future. Besides, glycoengineering strategies have been developed to reduce fucosylation or increase sialylation of biologies. Improving the glyco- sylation profile of biologies will definitively continue being a priority for the industry in order to enhance their quality and bioactivity.

**REFERENCES**


Lai, T., Yang, Y., Ng, S.K., 2013. Advances in Mammalian cell line development technologies for recombinant protein production. Pharmaceuticals (Basel) 6, 579–603.


