ELSEVIER

DRUG DISCOVERY TODAY TECHNOLOGIES

Drug Discovery Today: Technologies

Vol. 38, 2020



Biopharmaceuticals and Biosimilars

# Recent advances in CHO cell line development for recombinant protein production

Borbála Tihanyi, László Nyitray*

Department of Biochemistry, Eötvös Loránd University, Pázmány Péter stny 1/C, 1117 Budapest, Hungary


**Recombinant proteins used in biomedical research, diagnostics and different therapies are mostly produced in Chinese hamster ovary cells in the pharmaceutical industry. These biotherapeutics, monoclonal antibodies in particular, have shown remarkable market growth in the past few decades. The increasing demand for high amounts of biologics requires continuous optimization and improvement of production technologies. Research aims at discovering better means and methods for reaching higher volumetric capacity, while maintaining stable product quality. An increasing number of complex novel protein therapeutics, such as viral antigens, vaccines, bi- and tri-specific monoclonal antibodies, are currently entering industrial production pipelines. These biomolecules are, in many cases, difficult to express and require tailored product-specific solutions to improve their transient or stable production. All these requirements boost the development of more efficient expression optimization systems and high-throughput screening platforms to facilitate the design of product-specific cell line engineering and production strategies. In this mini-review, we provide an overview on recent advances in CHO cell line development, targeted genome manipulation techniques, selection systems and screening methods currently used in recombinant protein production.**

## Introduction

Biological therapeutics and diagnostics (theranostics) are rapidly growing products in the pharmaceutical market. They include various monoclonal antibodies (mAbs), vaccines, hormones, and other proteins, all of which have wide range of applications. Since 2002 there have been more than 300 biopharmaceutical approvals by FDA (U.S. Food and Drug Administration), and this number is growing, as biologics (proteins, nucleic acids, sugars and their complexes) are entering more and more fields in diagnostics and therapies [1]. Meeting the increasing demand for these products remains a challenge and drives constant innovation in the manufacturing processes. Strategies aim at optimizing protein expression to achieve higher volumetric productivity along with stable product quality and low manufacturing costs, in reduced times.

A significant portion of available biologics are recombinant proteins, most of which are produced in mammalian expression platforms. In this review, we focus on this expression system, although other systems for producing active recombinant proteins are also available and under evaluation for potential use in the pharmaceutical industry. Mammalian

*Corresponding author.: L. Nyitray (nyitray@elte.hu)

 https://doi.org/10.1016/j.ddtec.2021.02.003

cells tend to become predominant hosts for industrial biopharmaceutical production, largely because of their ability to produce diverse, correctly folded and glycosylated proteins. The importance of posttranslational modifications, glycosylation in particular, and the question of how they affect different properties of recombinant proteins has attracted much attention over the last decade. Extensive research led to the discovery that for potential therapeutic use large and complex proteins need to have human-like posttranslational modifications in order to be functional and non-immunogenic. Several studies from different groups have demonstrated that proper glycosylation profile promotes biological activity and stability, increases half-life and reduces immunogenicity of protein therapeutics [2–4].

## The CHO expression platform for pharmaceutical recombinant protein expression

Mammalian expression-based systems involve various cell lines of different origins, from hamster (CHO, BHK), human (HEK293, HT-1080, PER.C6, CAP, HKB-11, Huh-7) and mouse (NS0, Sp2/0). However, 70% of biologics, and almost all mAbs , are produced in Chinese hamster ovary (CHO) cells, as the most commonly used and preferred hosts for biopharmaceutical protein production [5]. Their popularity lies in the facts that they are capable of high productivity (0.1–1 g/L in batch and 1–10 g/L in fed-batch cultures), exhibit consistently good growth phenotypes, are suitable for large-scale industrial culturing, can easily be adapted to various chemically defined media, are less susceptible to infections by human viruses, and they are able to perform human compatible glycosylation [1,3,4,6].

CHO *ori* cell line was isolated and immortalized by Theodore Puck in 1956 [7]. Subclones were then further developed by introducing mutations or via adaptation to different culture conditions. CHO *ori*-derived cell lineages own specific properties favouring more efficient protein production or producer cell line establishment, as compared to the parental cell line. CHO DXB11 (*dhfr*+/−), CHO DG44 (*dhfr*−/−) and CHO-GS-/- (CHO-K1SV) hosts have been generated to make clone selection easier and faster by utilizing metabolic markers, the dihydrofolate reductase-mediated methotrexate (MTX)-based and the glutamine synthetase-mediated methionine sulfoximine (MSX)-based selection systems, respectively. At the same time, MTX selection also works as a gene amplification system, which allows for high product yields [3,4]. CHO-S and derived cell lines have been adapted to suspension culture with the aim to achieve higher volumetric production by increasing producer cell density.

Despite sharing the same common ancestor, different CHO lineages exhibit substantial genetic heterogeneity as a result of extensive mutagenesis and clonal selection [8,9]. These genetic differences, which affect karyogram and epigenetic variations, as well, account for host cell-specific differences among the most widely used CHO cell lines (mostly CHO-K1, CHO DXB11, CHO-S and CHO DG44). Thus, CHO expression platform offers a wide variety of choices for host cell selection and protein production optimization. Comparative studies from different groups provide large amount of data on cell-specific growth and product formation of different recombinant CHO cell lines for various biologics [10]. In a representative example, Reinhart et al. have shown that CHO-K1 cells favoured cell-specific productivity (7–16 pg/cell/day), wheres CHO-S had a preference for biomass production, but lower mAb expression (2–6 pg/cell/day), which was similar in CHO DG44 cells ((2–4 pg/cell/day) in the same experimental setup. The group also investigated the preference of CHO hosts for different culture conditions (batch, fed-batch, semi-continuous perfusion culture) and chemically defined media. The study concluded that the observed differences in expression were related to host cell-specific phenotype and metabolism, and showed no strong correlation with the cultivation method or cell culture media [10].

## CHO cell line optimization – a historical perspective

For decades, research groups all over the world have put tremendous efforts into trying to improve the performance of producer CHO cell lines. These attempts for cell line manipulation aimed to achieve higher productivity and product yields and took advantage of the pre-existing knowledge on transcription, translation, cell metabolism, signalling pathways and secretion machinery. A main strategy for host cell engineering was based on overexpressing genes beneficial for cell proliferation, longevity, stress and apoptosis resistence, protein production and secretion. Numerous studies have demonstrated that transient or stable overexpression of key genes involved in cellular metabolism, protein biosynthesis and glycosylation increased growth rates, productivity, and resulted in better product quality, respectively [11]. Cell viability and culture performance have also been shown to be improved by the overexpression of various transcription factors, anti-apoptotic (e.g. BCL2, XIAP, AVEN, MCL1) and pro-proliferative genes (reviewed by Fischer et al. [11]). Advances in proteomics have also contributed to the rational engineering of host cells. Analysis of proteomics data was used to identify potential targets for gene manipulation in order to enhance product yields, as reported for the glutamate-cysteine ligase modifier subunit [12].

In contrast to the forced expression of beneficial genes, other approaches for cell line optimization focused on the elimination or repression of' disadvantageous' genes by genomic knock-out or via RNAi-mediated gene silencing. There are different ways for disrupting gene function or eliminating gene products: i) partial or complete deletion of a specific gene by targeted genome manipulation, ii) switching off'' gene function by mutagenesis, iii) siRNA-mediated gene silencing. MicroRNAs, which are known to play an essential

role in transcriptome regulation in mammalian cells, have also been intensively investigated in the past decade for their potential in CHO cell engineering. A growing number of evidence suggest that miRNAs have influence on several cellular functions such as transcription, translation, ribosome biogenesis, secretion, cell cycle, metabolism and cell death. Exploiting the miRNA regulation system thus represents a highly efficient molecular tool for the engineering of entire cellular pathways. miRNA overexpression (e.g. miR-2861, miR-23, miR-17) or knock-down by antagomiRs, or via viral vectors, sponge and tough decoy vectors (e.g. miR-7, miR-106b, miR-14 and many others) has been successfully used in CHO cell line optimization to adapt cells to stressful environment, temperature changes and to enhance protein production [13–17]. A recent study has reported the deletion of miR-744 using the CRISPR/Cas9 genome editing system with the aim of improving CHO bioprocess performance [18]. Manipulation of miRNA-mediated regulation have also been applied to boost the production of artifically designed difficult-to-express new biotherapeutics [9,19].

Gene ablation studies have historically been directed against metabolic (e.g. LDHA), pro-apoptotic (BAX, BAK) and anti-proliferative genes, and cell cycle checkpoint kinases (ATR). As expected, the elimination or silencing of these selected genes has resulted in improved culture performance, apoptosis resistance and enhanced product yields. Genes involved in chromatin remodeling (e.g. HDACs) and glycosylation (e.g. FUT8, SLC35C1) have also been targets for knock-out studies in order to get a better understanding of how cell lines can be manipulated to express recombinant proteins with tailored posttranslational modification profile [11].

All these achievements in host cell engineering are rightly considered as a success story, as they have led to the generation of CHO host derivatives outperforming their parental cell lines regarding growth, stress resistance and productivity. Engineered production cell lines have been shown to be able to achieve record high 3–7 g/L volumetric productivities [20]. However, data from monitoring production efficiency in the industrial manufacturing suggest that, in many cases, it is not possible to recover the full amounts of products from bioreactors in high-titer production processes. Due to this downstream processing bottleneck, products do not seem to benefit from very high titers and efforts to further increase titers likely yield diminishing returns. Furthermore, in many cases, recombinant cell lines have been observed to show decline in viability and productivity during long term cultivation, thus making the industrial community face further challenges.

## New focus areas in industrial cell line development

Before 2010, many of the biopharmaceutical companies followed the policy of developing their own production cell lines, but the past decade has seen a change in this approach. Now companies tend to use the same host cell line for production and for initial clinical trials as they have realized that this strategy simplifies the cycle of cell line development and optimizes the way to the clinic. Besides, it reduces the risk of product comparability issues and help manufacturers meet product quality requirements and consistency expectations [20].

These changing priorities have contributed to new directions in innovation regarding the CHO expression platform. The focus of industrial cell line development has shifted towards the issues of host cell line stability for ensuring stable long term production, the use of targeted integration techniques for expression optimization, especially for complex engineered recombinant therapeutics, and the development of selection and screening systems for faster and more effective clone selection (Fig. 1).

### *CHO cell line instability*

All CHO cells, used in the pharmaceutical industry, are characterized by good adaptability to genetic manipulations and changing culture conditions. This feature is the consequence of the inherently plastic genome of these cell lines and it makes them amenable for large-scale industrial production [1]. On the downside, this cell plasticity gives CHO cells higher propensity for genomic rearrangements (deletions, translocations) and it becomes a source for cell line instability during the bioproduction process. Producer cells are under high genomic and metabolic demand, which may cause genomic and epigenetic changes resulting in a decline in productivity and product quality [21,22]. Despite all this, the numerous advantages offered by CHO cell lines make them ideal hosts for recombinant protein production, thus, it is unlikely that the scientific and industrial community will abandon this expression system in the near future. Therefore, much effort has recently been put into understanding the phenomenon of CHO cell line instability. Research focuses on finding solutions to overcome the problem of phenotypic instability during the manufacturing process to ensure long term stable production, process consistency and to assure acceptable product quality.

CHO host cell lines, as all other rapidly growing immortalized cells, are genomically unstable and intraclonally heterogenous, which gives the system robustness and flexibility, but causes problems with production stability. To better understand the heterogenous nature of CHO cells, Borth and her group investigated in a series of experiments how heterogeneity is transferred during subcloning and clonal selection for producer cell line generation. They found that subcloning itself did not result in reduced genetic and phenotypic heterogeneity in subclones as compared to the original cell pool, either in host or in recombinant cell lines. On the contrary, selection for a specific cell property led to increased karyotypic homogeneity and chromosomal stabili-

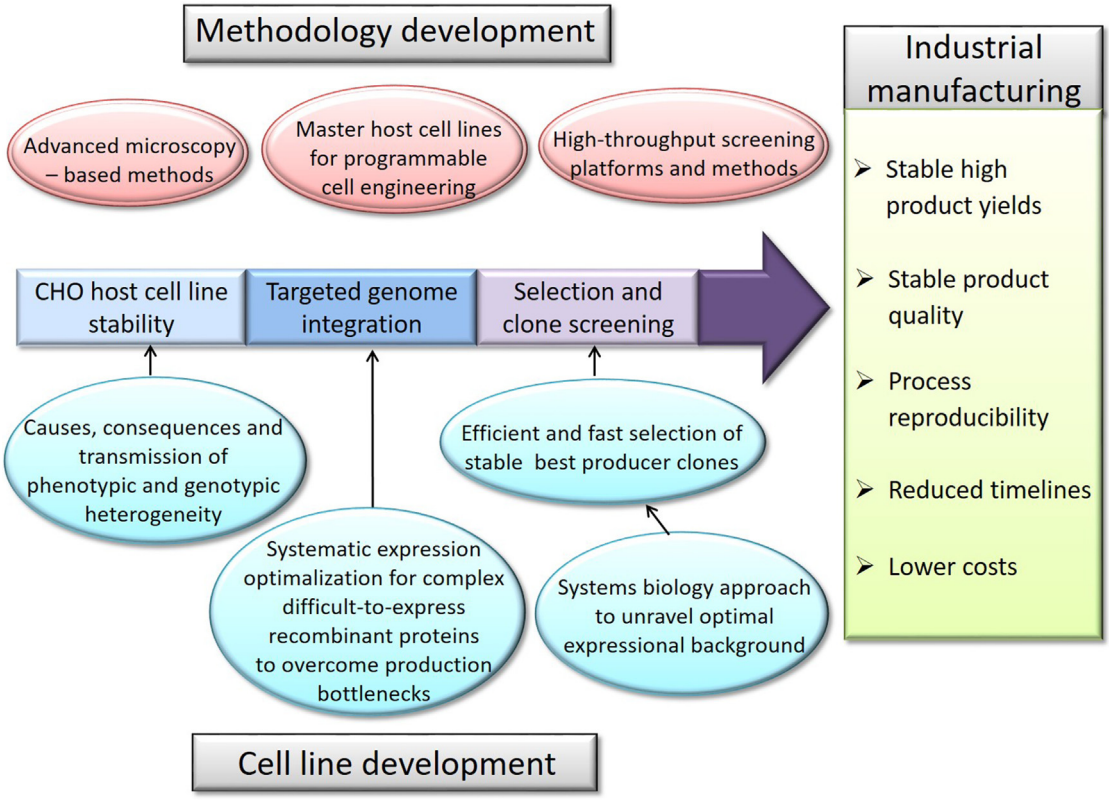


**Fig. 1.** Current focus areas of innovative research in industrial cell line development.

ty [23]. The connection among CHO cell line instability, karyotype variation and production instability has also been investigated by Baik and Lee. They proposed a model for the mechanism of production instability, where random chromosomal rearrangements were associated with growth advantage and non- or low-producing clones. At the same time, MTX selection for a specific phenotype/genotype reduced growth rates but improved product yields [24,25].

An important issue during the development of stable recombinant cell lines is to generate and identify high producing clones, which do not lose their expression capability over time. (Clonality is considered critical in therapeutic protein production, as clonal lineages are thought to offer a stable and consistent product quality profile.) Production of selected clones then needs to be assessed during long term cultivation, which is a labour intensive and time consuming process. However, the rapidly accumulating transcriptomic data available on CHO cell lines open the way for functional analysis to identify regulators and biomarkers of high production and clonal stability [26,27]. Following this approach, Ritter and co-workers have analyzed the gene expression profiles of low and high producing CHO clones to find potential targets for cell engineering to make clone selection faster and more efficient. They found that the majority of stable high producing clones were marked by the loss of the telomeric region of chromosome 8. Results from the established experimental model confirmed that CHO-C8DEL cells (with the telomeric region of chromosome 8 deleted) exhibited higher volumetric productivity and generated significantly more stable producer clones, as compared to control CHO-K1 cells [28]. Further on, the group performed a series of knock-out and knock-down studies and identified two specific targets within the selected telomeric region, *Fam60A* and *C12orf35*, which played a role in creating the observed phenotype of CHO-C8DEL cells. Deletion or reduced expression of these genes was reported to correlate with higher productivity and faster recovery from selection [29,30].

### *Targeted genome manipulation techniques*

In recombinant cell line generation, the most intensive research line focuses on the use of targeted genome integration techniques for expression optimization. For the establishment of stable and clonal (isogenic) producer cell lines, transgenes need to be integrated into the host genome. The conventional method for this relied on the random integration of expression vectors, followed by a time consuming screening for suitable cells from a recombinant cell pool with heterogenous transgene insertions. This approach has two major shortcomings: the low efficiency of integration and the position effect at the integration sites. There are

different attempts to overcome these problems. Integration efficiency was successfully improved with the help of different transposon systems (Sleeping Beauty, Leap-in, PiggyBac, Tol2). Recombinant CHO pools generated in only 2–3 weeks using transposon-mediated gene integration have been reported to be able to achieve extraordinarily high product titers (>7 g/L) and hence can be used for rapid production of large amounts of proteins [31,32]. Gene integration with the help of transposon systems is still an area of interest and is under evaluation for its potential use in different areas of biomedical research.

The problem of 'position effect' occurring as a result of random integration has also been addressed by several research groups over the past decade. Ubiquitous Chromatin-opening Elements (UCOEs), reviewed by Neville et al., and other regulatory motifs identified in the genome of high producer cells have been applied to prevent gene silencing and to maintain high levels of expression [33,34]. Bacterial Artificial Chromosomes (BACs) also represent a popular and widely used alternative method for ensuring site-independent, stable and high-level expression for integrated transgenes (e.g. 0.75 g/L in lab-scale bioreactors and 1.12 g/L in shake flasks for selected antibodies) [10,35].

Targeted genome manipulation techniques enable the knock-in of recombinant protein coding genes into well-defined and transcriptionally active genomic sites. This is a much faster and far more efficient way for establishing producer clones, as compared to random integration. Targeted genetic modification of cells has been made possible by the emergence of target-specific genome-editing tools, such as zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), clustered regularly interspaced short palindromic repeats (CRISPR)-CRISPR associated protein 9 (Cas9) RNA guided nucleases [36–38]. Among these, the CRISPR/Cas9 system became the most popular for its ease of use (as compared to ZFNs and TALENs), high editing efficiency and low cost. Programmable nucleases generate double strand breaks, which are resolved in mammalian cells by endogenous DNA repair mechanisms. These comprise different pathways: i) non-homologous end joining (NHEJ), ii) homology-directed repair (HR) and iii) microhomology-mediated end-joining (MMEJ). The main applications of targeted genome editing tools involve the knock-out or knock-in of genes of interest in the genome of cells or organisms.

Much effort has been put into identifying transcriptional' hot spots' in host cell genomes in recent years. Target sites have been chosen based on previous empirical data available, transcriptomic data analysis or with the help of lentiviral screens [39]. Locuses used recently by different research groups involve of *Fer1L4, Rosa26, Hprt, Hipp11, C12orf35 and GS* [39–42,72–74]. Today, the state-of-art technique is to integrate not only individual transgenes, but also complex cassettes, called' landing pads', to generate master host cell lines. Landing pads contain selection markers and recognition sites for recombinases or integrases, which allow the precise insertion of any kind and any number of expression cassettes into the established sites. There are numerous systems that have been effectively used for mediating expression cassette exchange by site-specific recombination in mammalian cells, such as Cre-loxP [40,72], Flp-FRT [41], BxB1 recombinase [39,41], PhiC31 integrase [42] (see Table 1).

A recent study has reported the development of multi-landing pad DNA integration platforms, with 3–4 landing pads inserted into different genomic hot spots, for advanced cell engineering [39]. Alternative versions, by other groups, have been generated to allow the cumulative integration of multiple copies of the same gene or different genes [40]. Such constructions may bring about the problem of repeat-mediated gene silencing (also affecting polycistronic constructions and genes in close vicinity of each other), however, various methods are available (e.g. the use of epigenetic modifying elements, such as insulators, UCOEs) to prevent the silencing of transgenes in CHO cells [43,44]. These platforms discussed above have been developed to facilitate the generation of recombinant producer cell lines in the pharmaceutical industry. They are also valuable and promising tools for the expression optimization of novel complex protein therapeutics and might be applied in synthetic biology, as well.

### *Expression optimization for difficult-to-express recombinant protein therapeutics with regard to product quality and quantity*

A large number of engineered, complex recombinant proteins are currently filling industrial production pipelines. Despite using advanced CHO cell expression platforms, these

**Table 1. Targeted genomic hotspots and methods for expression construct integration and design in CHO cell lines.**

| Transcriptionally active genomic locus for transgene or' landing pad' integration | Method for site specific expression cassette exchange | Reference |
|---|---|---|
| *Fer1L4 (CRISPR/Cas9)* | Bxb1 recombinase Flp-FRT | [41] |
| *Rosa26 (CRISPR/Cas9)* | Bxb1 recombinase | [39] |
| *hprt (CRISPR/Cas9)* | Cre-loxP | [40,72] |
| *Hipp11 (CRISPR/Cas9)* | PhiC31 integrase | [42] |
| *C12orf35, hprt, GRIK1 (CRISPR/Cas9)* | | [73] |
| *GS (glutamine synthetase) (ZFN vs CRISPR/Cas9)* | | [74] |

proteins, in many cases, prove to be difficult-to-express'' (DTE). A major area of innovative research in the pharmaceutical industry focuses on improving the manufacturability of these bioproducts. There have been numerous attempts from different research groups to try to understand the causes for production bottlenecks and to seek solutions for these problems recently.

The first key question to be addressed is expression optimization at the level of transcription, which involves expression construct design, method of genomic integration, and interaction with the genomic environment of the integration site. Transcriptional levels of recombinant proteins even at well-characterized integration sites are affected by the interplay between the integration site and the expression cassette components. The investigation of this phenomenon is highly necessary for programmable cell engineering to be able to predict how multiple components contribute to the final net expression of inserted transgenes. To address this question, Kildegaard and co-workers have developed a molecular toolbox for the systematic evaluation of product and site specific recombinant gene expression. Using CRISPR/Cas9 technology, they constructed a series of CHO-S model cell lines with targeted integration site for a landing pad containing a recombinant gene under defined 5′ proximal regulatory elements. They showed that different regulatory elements generated robust recombinant gene expression patterns at defined integration sites with a wide range of transcriptional outputs [45]. Cartwright and co-workers have reported the development of a high-throughput microscale platform for the multiparallel testing of different genetic components and for the systematic assessment of their effects on DTE protein production. The group aimed to optimize cell engineering solutions for both stable and transient productions of a model DTE IgG. They concluded that optimal methods were product and genetic context specific and significantly differed for stable and transient productions [46]. Tadauchi et al. have conducted a systematic investigation aiming to understand what makes antibody expression difficult. They have constructed model cell lines with landing pads at defined genomic integration sites, where the inserted transgene was driven under the control of an (Tet/Dox) inducible promoter. This system made it possible to characterize the expression of proteins, which may be toxic or for any reason difficult-to-express [47]. Other works have focused on transcriptional regulation, as in the case of miRNA mediated (miR-557) engineering of host cells, which has been shown to help enhancing DTE protein expression [48].

Secretion deficiencies comprise another main field of interest in the case of complex therapeutics, especially DTE mAbs (e.g. infliximab) [49]. There are various approaches from different groups to face this problem. In recent papers, Hussein et al. have described a novel protein engineering strategy to bypass secretory bottlenecks and thus to improve production of DTE proteins. The group performed a step-by-step analysis to understand the molecular mechanisms which restricted the production of the selected recombinant proteins, TIMP-3 and TIMP-4, and identified a posttranslational processing block as a limiting step in protein expression. To overcome this limitation, a fusion protein was created that included a short pro-sequence from a secretory growth factor with a native cleavage site for the protease furin. At the same time, furin was overexpressed in the same experimental cell line. This strategy has been successfully used to improve protein production in the case of two members of the TIMP family which originally showed poor secretion. Following this study, a computational tool has been applied to analyze TIMP-3 coding nucleotide sequence to identify sequence features which contributed to poor secretion. Based on the data acquired, the group constructed a model for a protein chimera design to be used for DTE recombinant targets [50,51]. Otte and her group have reported the establishment of an automated confocal microscopy-based method for the investigation of DTE protein expression within the cells to explore the causes of production bottlenecks. In representative studies, Mathias et al. followed the processing of selected recombinant DTE proteins during cellular synthesis and investigated their distribution within the respective organelles of the secretory pathway. They showed that in the case of a selected antibody wrong folding was the rate limiting step and the reason for hampered secretion, as it caused the ER-associated degradation of the molecule [52,53].

Besides volumetric productivity, product quality is also of key importance for biotherapeutic protein production and is constantly in the focus of industrial technology development efforts. Glycoproteins are the fastest-growing class of biotherapeutics. Correct posttranslational modifications, especially glycan structures are crucial for the potency and control of pharmacokinetic and pharmacodynamic properties of these biologics. Moreover, it is not only the nature and amount of glycosylation (sialylation, fucosylation) what matters, but also the heterogeneity of N-glycans can be an issue. This heterogeneity, arising due to the variability of N-glycan processing, may compromise the activity and safety of glycotherapeutics [54]. Yang and co-workers have described how CHO cells can be genetically engineered to produce glycoproteins in a nearly homogenous form. They performed a knockout screen for CHO glycosyltransferases to identify key genes that control decisive steps in N-linked glycosylation and showed that CHO cells tolerated glycoengineering without compensatory changes. Based on their results, they provided a strategy for the reconstruction of more homogeneous glycosylation capacities in the model cell lines [55].

### *Recent advances in the development of high-throughput selection and screening systems*

Selection systems have become standardized during the decades, however, the advantages versus disadvantages of differ-

ent selection systems continue to form a matter of debate, especially in the case of the MTX selection system. Several studies have demonstrated that the gene amplification resulting from MTX selection is a major cause for production instability in recombinant cell lines. Moreover, multiple rounds of MTX selection require a longer time for cell line generation. Recently, Yeo et al. have also published a large-scale comparative work on how various types of selection marker genes affect production stability in different CHO cell lines. Their results indicate that selection marker expression needs to be optimized for different types of CHO cells to improve selection stringency for best producer clones [56].

With the growing demand for biotherapeutic products in the market, time and expenses are becoming important factors regarding the manufacturing process. During process design and optimization more attention is devoted to the problem of how to reduce the time and costs required for recombinant cell line generation for novel bioproducts. This has boosted the development of novel methods and techniques for high-throughput clone screening and characterization. Many improvement efforts focus on monocloning, as the most time consuming process during producer cell line establishment. Several new techniques are now available for more efficient single cell isolation replacing the classical method of ''limiting dilution''. Modern cloning approaches employ specialized instruments to ensure that a single cell is seeded into microtiter plate wells. Fluorescence-activated cell sorting (FACS) has been successively used for sorting single cells with specific cellular attributes indicative of high productivity. VIPS (Verified in-situ plate seeding), developed by Solentim, as well as Cytena single cell printer instruments combine cell seeding with microscopic imaging to assure single cell deposition, thereby the single cell origin of derived clones [57–59]. The Beacon-platform from Berkeley Lights which utilizes nanofluidics and optoelectro positioning technology allows the culturing, manipulation and characterization of thousands of cells in parallel on a single chip, via software-controlled operations [60,61]. ClonePX FL offers an expression based method, on semi solid medium, for efficient high producer clone / colony selection and isolation [62].

Many of these technology developments and modern instruments providing possibility for high-throughput automated processes have already been successfully introduced in the pharmaceutical industry and are now applied in industrial recombinant protein production.

### *'Omics technologies – a systems biology approach*

Since the 2010's, the emergence of 'omics approaches and 'omics technologies has contributed to the deeper understanding of cellular characteristics and identity profiles that are indicative of optimal production. A vast amount of data is now available (and is created day by day) in the field of genomics, transcriptomics, proteomics, metabolomics, lipidomics and glycomics, which provide a systems level insight into cellular states and characteristics in a given moment. Stolfa et al. have published a comprehensive summary on available resources and databases for CHO [63]. In parallel, the constant development of computational methods and algorithms in bioinformatics enables the translation'' of these 'omics data into biologically relevant information on cellular mechanisms and functions [64,65]. These information together with computational modeling are used in all fields of CHO research: for predicting cell behaviour [64], to explore transcriptome level clonal variation underlying different production efficiencies in recombinant cell lines [66], or to identify novel targets for manipulating stress response and sensitivity in producer cell lines [67]. In a recent study, Kol et al. have carried out multiplex secretome analysis to investigate the impact of host cell proteins on recombinant protein production [68]. By systematic *in silico* (computational screening) and *in vitro* analysis, Nguyen et al. have identified endogenous promoters and enhancer elements from CHO cells for potential use in transgene expression optimization [69]. Brown and co-workers have performed transcriptome-based analysis to identify optimal reference CHO genes for qPCR data normalisation. Results of the study have been adapted for routine use in the pharmaceutical industry for producer cell line stability characterization [70]. With the advances in 'omics technology, the system-wide analysis of glycoprofiles provides new insights into the glycosylation network of CHO cells and helps to develop novel glycoengineering strategies with the aim of generating more efficacious therapies with fewer side effects – as summarized by Tejwani et al. in a comprehensive review [54]. Stach et al. describes a method for combining system-level kinetic models with synthetic biology to investigate the nature of N-glycosylation and to test the synergistic behavior of genetic perturbations on the galactosylation of a produced IgG [71].

## Future perspectives

Systems biology approach has opened a new dimension for enhancing CHO-based bioproduction.' Omics technologies provide new opportunities to better characterize and understand complex cellular functions, thus create a basis for the possibility of mechanism-driven cell line optimization, instead of manipulating a few key genes. However, a main problem is that these modern technologies require special instruments, materials, advanced information technology facility, therefore they are still too complex and expensive to be routinely used in industrial manufacturing. On the other hand, data provided by different 'omics technologies so far cannot be integrated into a general mathematical model, which would facilitate the transfer and application of these information in experimental process design. How this knowledge can be adapted and integrated into industrial bioprocess optimization and development regarding recom-

binant protein production still remains a challenge and needs to be addressed in the future.

## Acknowledgement

The authors acknowledge the support from the National Research, Development and Innovation Office, funded under the ''Competitiveness and excellence cooperations'' (2017-1.3.1-VKE-201700002), and the ''National Excellence'' programmes (2018-1.2.1-NKP-2018-00005 and TKP2020-IKA-05).

## References

[1] Dahodwala H, Lee KH. The fickle CHO: a review of the causes, implications, and potential alleviation of the CHO cell line instability problem. Curr Opin Biotechnol 2019;60:128–37. http://dx.doi.org/10.1016/j.copbio.2019.01.011.

[2] Kuriakose A, Chirmule N, Nair P. Immunogenicity of biotherapeutics: causes and association with posttranslational modifications. J Immunol Res 2016;20161298473. http://dx.doi.org/10.1155/2016/1298473.

[3] Kim JY, Kim YG, Lee GM. CHO cells in biotechnology for production of recombinant proteins: current state and further potential. Appl Microbiol Biotechnol 2012;93(3):917–30. http://dx.doi.org/10.1007/s00253-011-3758-5.

[4] Lai T, Yang Y, Ng SK. Advances in Mammalian cell line development technologies for recombinant protein production. Pharmaceuticals (Basel) 2013;6(5):579–603. http://dx.doi.org/10.3390/ph6050579.

[5] Lalonde ME, Durocher Y. Therapeutic glycoprotein production in mammalian cells. J Biotechnol 2017;251:128–40. http://dx.doi.org/10.1016/j.jbiotec.2017.04.028.

[6] Kunert R, Reinhart D. Advances in recombinant antibody manufacturing. Appl Microbiol Biotechnol 2016;100(8):3451–61. http://dx.doi.org/10.1007/s00253-016-7388-9.

[7] Puck TT, Cieciura SJ, Robinson A. Genetics of somatic mammalian cells. III. Long-term cultivation of euploid cells from human and animal subjects. J Exp Med 1958;108(6):945–56. http://dx.doi.org/10.1084/jem.108.6.945.

[8] Bailey LA, Hatton D, Field R, Dickson AJ. Determination of Chinese hamster ovary cell line stability and recombinant antibody expression during long-term culture. Biotechnol Bioeng 2012;109(8):2093–103. http://dx.doi.org/10.1002/bit.24485.

[9] Tharmalingam T, Barkhordarian H, Tejeda N, Daris K, Yaghmour S, Yam P, et al. Characterization of phenotypic and genotypic diversity in subclones derived from a clonal cell line. Biotechnol Prog 2018;34(3):613–23. http://dx.doi.org/10.1002/btpr.2666.

[10] Reinhart D, Damjanovic L, Kaisermayer C, Sommeregger W, Gili A, Gasselhuber B, et al. Bioprocessing of recombinant CHO-K1, CHO-DG44, and CHO-S: CHO expression hosts favor either mAb production or biomass synthesis. Biotechnol J 2019;14(3)e1700686. http://dx.doi.org/10.1002/biot.201700686.

[11] Fischer S, Handrick R, Otte K. The art of CHO cell engineering: a comprehensive retrospect and future perspectives. Biotechnol Adv 2015;33(8):1878–96. http://dx.doi.org/10.1016/j.biotechadv.2015.10.015.

[12] Orellana CA, Marcellin E, Gray PP, Nielsen LK. Overexpression of the regulatory subunit of glutamate-cysteine ligase enhances monoclonal antibody production in CHO cells. Biotechnol Bioeng 2017;114(8):1825–36. http://dx.doi.org/10.1002/bit.26316.

[13] Fischer S, Paul AJ, Wagner A, Mathias S, Geiss M, Schandock F, et al. miR-2861 as novel HDAC5 inhibitor in CHO cells enhances productivity while maintaining product quality. Biotechnol Bioeng 2015;112(10):2142–53. http://dx.doi.org/10.1002/bit.25626.

[14] Jadhav V, Hackl M, Klanert G, Hernandez Bort JA, Kunert R, Grillari J, et al. Stable overexpression of miR-17 enhances recombinant protein production of CHO cells. J Biotechnol 2014;175:38–44. http://dx.doi.org/10.1016/j.jbiotec.2014.01.032.

[15] Kelly PS, Breen L, Gallagher C, Kelly S, Henry M, Lao NT, et al. Re-programming CHO cell metabolism using miR-23 tips the balance towards a highly productive phenotype. Biotechnol J 2015;10(7):1029–40. http://dx.doi.org/10.1002/biot.201500101.

[16] Coleman O, Suda S, Meiller J, Henry M, Riedl M, Barron N, et al. Increased growth rate and productivity following stable depletion of miR-7 in a mAb producing CHO cell line causes an increase in proteins associated with the Akt pathway and ribosome biogenesis. J Proteomics 2019;195:23–32. http://dx.doi.org/10.1016/j.jprot.2019.01.003.

[17] Xu C, Han Q, Zhou Q, Zhang L, Wu P, Lu Y, et al. MiR-106b promotes therapeutic antibody expression in CHO cells by targeting deubiquitinase CYLD. Appl Microbiol Biotechnol 2019;103(17):7085–95. http://dx.doi.org/10.1007/s00253-019-10000-3.

[18] Raab N, Mathias S, Alt K, Handrick R, Fischer S, Schmieder V, et al. CRISPR/Cas9-mediated knockout of MicroRNA-744 improves antibody titer of CHO production cell lines. Biotechnol J 2019;14(5)e1800477. http://dx.doi.org/10.1002/biot.201800477.

[19] Schoellhorn M, Fischer S, Wagner A, Handrick R, Otte K. miR-143 targets MAPK7 in CHO cells and induces a hyperproductive phenotype to enhance production of difficult-to-express proteins. Biotechnol Prog 2017;33(4):1046–58. http://dx.doi.org/10.1002/btpr.2475.

[20] Kelley B, Kiss R, Laird M. A different perspective: how much innovation is really needed for monoclonal antibody production using mammalian cell technology? Adv Biochem Eng Biotechnol 2018;165:443–62. http://dx.doi.org/10.1007/10_2018_59.

[21] Feichtinger J, Hernández I, Fischer C, Hanscho M, Auer N, Hackl M, et al. Comprehensive genome and epigenome characterization of CHO cells in response to evolutionary pressures and over time. Biotechnol Bioeng 2016;113(10):2241–53. http://dx.doi.org/10.1002/bit.25990.

[22] Li H, Chen K, Wang Z, Li D, Lin J, Yu C, et al. Genetic analysis of the clonal stability of Chinese hamster ovary cells for recombinant protein production. Mol Biosyst 2016;12(1):102–9. http://dx.doi.org/10.1039/c5mb00627a.

[23] Vcelar S, Melcher M, Auer N, Hrdina A, Puklowski A, Leisch F, et al. Changes in chromosome counts and patterns in CHO cell lines upon generation of recombinant cell lines and subcloning. Biotechnol J 2018;13(3)e1700495. http://dx.doi.org/10.1002/biot.201700495.

[24] Baik JY, Lee KH. A framework to quantify karyotype variation associated with CHO cell line instability at a single-cell level. Biotechnol Bioeng 2017;114(5):1045–53. http://dx.doi.org/10.1002/bit.26231.

[25] Noh SM, Shin S, Lee GM. Comprehensive characterization of glutamine synthetase-mediated selection for the establishment of recombinant CHO cells producing monoclonal antibodies. Sci Rep 2018;8(1):5361. http://dx.doi.org/10.1038/s41598-018-23720-9.

[26] Datta P, Linhardt RJ, Sharfstein ST. An' omics approach towards CHO cell engineering. Biotechnol Bioeng 2013;110(5):1255–71. http://dx.doi.org/10.1002/bit.24841.

[27] Kildegaard HF, Baycin-Hizal D, Lewis NE, Betenbaugh MJ. The emerging CHO systems biology era: harnessing the' omics revolution for biotechnology. Curr Opin Biotechnol 2013;24(6):1102–7. http://dx.doi.org/10.1016/j.copbio.2013.02.007.

[28] Ritter A, Voedisch B, Wienberg J, Wilms B, Geisse S, Jostock T, et al. Deletion of a telomeric region on chromosome 8 correlates with higher productivity and stability of CHO cell lines. Biotechnol Bioeng 2016;113(5):1084–93. http://dx.doi.org/10.1002/bit.25876.

[29] Ritter A, Nuciforo S, Schulze A, Oertli M, Rauschert T, Voedisch B, et al. Fam60A plays a role for production stabilities of recombinant CHO cell lines. Biotechnol Bioeng 2017;114(3):701–4. http://dx.doi.org/10.1002/bit.26181.

[30] Ritter A, Rauschert T, Oertli M, Piehlmaier D, Mantas P, Kuntzelmann G, et al. Disruption of the gene C12orf35 leads to increased productivities in recombinant CHO cell lines. Biotechnol Bioeng 2016;113(11):2433–42. http://dx.doi.org/10.1002/bit.26009.

[31] Rajendra Y, Balasubramanian S, Peery RB, Swartling JR, McCracken NA, Norris DL, et al. Bioreactor scale up and protein product quality characterization of piggyBac transposon derived CHO pools. Biotechnol Prog 2017;33(2):534–40. http://dx.doi.org/10.1002/btpr.2447.

[32] Balasubramanian S, Peery RB, Minshull J, Lee M, White R, Kelly RM, et al. Generation of high expressing Chinese hamster ovary cell pools using the leap-in transposon system. Biotechnol J 2018;13(10)e1700748. http://dx.doi.org/10.1002/biot.201700748.

[33] Neville JJ, Orlando J, Mann K, McCloskey B, Antoniou MN. Ubiquitous Chromatin-opening Elements (UCOEs): applications in biomanufacturing and gene therapy. Biotechnol Adv 2017;35(5):557–64. http://dx.doi.org/10.1016/j.biotechadv.2017.05.004.
[34] Takagi Y, Yamazaki T, Masuda K, Nishii S, Kawakami B, Omasa T. Identification of regulatory motifs in the CHO genome for stable monoclonal antibody production. Cytotechnology 2017;69(3):451–60. http://dx.doi.org/10.1007/s10616-016-0017-8.
[35] Zboray K, Sommeregger W, Bogner E, Gili A, Sterovsky T, Fauland K, et al. Heterologous protein production using euchromatin-containing expression vectors in mammalian cells. Nucleic Acids Res 2015;43(16): e102. http://dx.doi.org/10.1093/nar/gkv475.
[36] Lee JS, Grav LM, Lewis NE, Faustrup Kildegaard H. CRISPR/Cas9-mediated genome engineering of CHO cell factories: application and perspectives. Biotechnol J 2015;10(7):979–94. http://dx.doi.org/10.1002/biot.201500082.
[37] Kim H, Kim JS. A guide to genome engineering with programmable nucleases. Nat Rev Genet 2014;15(5):321–34. http://dx.doi.org/10.1038/nrg3686.
[38] Sakuma T, Takenaga M, Kawabe Y, Nakamura T, Kamihira M, Yamamoto T. Homologous recombination-independent large gene cassette knock-in in CHO cells using TALEN and MMEJ-directed donor plasmids. Int J Mol Sci 2015;16(10):23849–66. http://dx.doi.org/10.3390/ijms161023849.
[39] Gaidukov L, Wroblewska L, Teague B, Nelson T, Zhang X, Liu J, et al. A multi-landing pad DNA integration platform for mammalian cell engineering. Nucleic Acids Res 2018;46(8):4072–86. http://dx.doi.org/10.1093/nar/gky216.
[40] Kawabe Y, Inao T, Komatsu S, Huang G, Ito A, Omasa T, et al. Improved recombinant antibody production by CHO cells using a production enhancer DNA element with repeated transgene integration at a predetermined chromosomal site. J Biosci Bioeng 2017;123(3):390–7. http://dx.doi.org/10.1016/j.jbiosc.2016.10.011.
[41] Inniss MC, Bandara K, Jusiak B, Lu TK, Weiss R, Wroblewska L, et al. A novel Bxb1 integrase RMCE system for high fidelity site-specific integration of mAb expression cassette in CHO cells. Biotechnol Bioeng 2017;114(8):1837–46. http://dx.doi.org/10.1002/bit.26268.
[42] Chi X, Zheng Q, Jiang R, Chen-Tsai RY, Kong LJ. A system for site-specific integration of transgenes in mammalian cells. PLoS One 2019;14(7) e0219842. http://dx.doi.org/10.1371/journal.pone.0219842.
[43] Saunders F, Sweeney B, Antoniou MN, Stephens P, Cain K. Chromatin function modifying elements in an industrial antibody production platform–comparison of UCOE, MAR, STAR and cHS4 elements. PLoS One 2015;10(4)e0120096. http://dx.doi.org/10.1371/journal.pone.0120096.
[44] Romanova N, Noll T. Engineered and natural promoters and chromatin-modifying elements for recombinant protein expression in CHO cells. Biotechnol J 2018;13(3)e1700232. http://dx.doi.org/10.1002/biot.201700232.
[45] Pristovsek N, Nallapareddy S, Grav LM, Hefzi H, Lewis NE, Rugbjerg P, et al. Systematic evaluation of site-specific recombinant gene expression for programmable mammalian cell engineering. ACS Synth Biol 2019;8 (4):758–74. http://dx.doi.org/10.1021/acssynbio.8b00453.
[46] Cartwright JF, Arnall CL, Patel YD, Barber NOW, Lovelady CS, Rosignoli G, et al. A platform for context-specific genetic engineering of recombinant protein production by CHO cells. J Biotechnol 2020;312:11–22. http://dx.doi.org/10.1016/j.jbiotec.2020.02.012.
[47] Tadauchi T, Lam C, Liu L, Zhou Y, Tang D, Louie S, et al. Utilizing a regulated targeted integration cell line development approach to systematically investigate what makes an antibody difficult to express. Biotechnol Prog 2019;35(2)e2772. http://dx.doi.org/10.1002/btpr.2772.
[48] Fischer S, Marquart KF, Pieper LA, Fieder J, Gamer M, Gorr I, et al. miRNA engineering of CHO cells facilitates production of difficult-to-express proteins and increases success in cell line development. Biotechnol Bioeng 2017;114(7):1495–510. http://dx.doi.org/10.1002/bit.26280.
[49] Kaneyoshi K, Kuroda K, Uchiyama K, Onitsuka M, Yamano-Adachi N, Koga Y, et al. Secretion analysis of intracellular "difficult-to-express" immunoglobulin G (IgG) in Chinese hamster ovary (CHO) cells. Cytotechnology 2019;71(1):305–16. http://dx.doi.org/10.1007/s10616-018-0286-5.
[50] Hussain H, Fisher DI, Abbott WM, Roth RG, Dickson AJ. Use of a protein engineering strategy to overcome limitations in the production of "Difficult to Express" recombinant proteins. Biotechnol Bioeng 2017;114 (10):2348–59. http://dx.doi.org/10.1002/bit.26358.
[51] Hussain H, Fisher DI, Roth RG, Abbott WM, Carballo-Amador MA, Warwicker J, et al. A protein chimera strategy supports production of a model "difficult-to-express" recombinant target. FEBS Lett 2018;592 (14):2499–511. http://dx.doi.org/10.1002/1873-3468.13170.
[52] Mathias S, Wippermann A, Raab A, Zeh N, Handrick R, Gorr I, et al. Unraveling what makes a monoclonal antibody difficult-to-express: from intracellular accumulation to incomplete folding and degradation via ERAD. Biotechnol Bioeng 2020;117(1):5–16. http://dx.doi.org/10.1002/bit.27196.
[53] Mathias S, et al. Visualisation of intracellular production bottlenecks in suspension-adapted CHO cells producing complex biopharmaceuticals using fluorescence microscopy. J Biotechnol 2018;271:47–55. http://dx.doi.org/10.1016/j.jbiotec.2018.02.009.
[54] Tejwani V, Andersen MR, Nam JH, Sharfstein ST. Glycoengineering in CHO cells: advances in systems biology. Biotechnol J 2018;13(3) e1700234. http://dx.doi.org/10.1002/biot.201700234.
[55] Yang Z, Wang S, Halim A, Schulz MA, Frodin M, Rahman SH, et al. Engineered CHO cells for production of diverse, homogeneous glycoproteins. Nat Biotechnol 2015;33(8):842–4. http://dx.doi.org/10.1038/nbt.3280.
[56] Yeo JHM, Ho SCL, Mariati M, Koh E, Tay SJ, Woen S, et al. Optimized selection marker and CHO host cell combinations for generating high monoclonal antibody producing cell lines. Biotechnol J 2017;12(12). http://dx.doi.org/10.1002/biot.201700175.
[57] Fieder J, Schulz P, Gorr I, Bradl H, Wenger T. A single-step FACS sorting strategy in conjunction with fluorescent vital dye imaging efficiently assures clonality of biopharmaceutical production cell lines. Biotechnol J 2017;12(6). http://dx.doi.org/10.1002/biot.201700002.
[58] Pekle E, Smith A, Rosignoli G, Sellick C, Smales CM, Pearce C. Application of imaging flow cytometry for the characterization of intracellular attributes in Chinese hamster ovary cell lines at the single-cell level. Biotechnol J 2019;14(7)e1800675. http://dx.doi.org/10.1002/biot.201800675.
[59] Yim M, Shaw D. Achieving greater efficiency and higher confidence in single-cell cloning by combining cell printing and plate imaging technologies. Biotechnol Prog 2018;34(6):1454–9. http://dx.doi.org/10.1002/btpr.2698.
[60] Le K, Tan C, Le H, Tan J, Zasadzinska E, Diep J, et al. Assuring clonality on the beacon digital cell line development platform. Biotechnol J 2020;15 (1)e1900247. http://dx.doi.org/10.1002/biot.201900247.
[61] Le K, Tan C, Gupta S, Guhan T, Barkhordarian H, Lull J, et al. A novel mammalian cell line development platform utilizing nanofluidics and optoelectro positioning technology. Biotechnol Prog 2018;34(6):1438–46. http://dx.doi.org/10.1002/btpr.2690.
[62] Roy G, Miro-Quesada G, Zhuang L, Martin T, Zhu J, Wu H, et al. Sequential screening by ClonePix FL and intracellular staining facilitate isolation of high producer cell lines for monoclonal antibody manufacturing. J Immunol Methods 2017;451:100–10. http://dx.doi.org/10.1016/j.jim.2017.08.012.
[63] Stolfa G, Smonskey MT, Boniface R, Hachmann A-B, Gulde P, Joshi AD, et al. CHO-omics review: the impact of current and emerging technologies on Chinese hamster ovary based bioproduction. Biotechnol J 2018;13(3) e1700227. http://dx.doi.org/10.1002/biot.201700227.
[64] Kyriakopoulos S, Ang KS, Lakshmanan M, Huang Z, Yoon S, Gunawan R, et al. Kinetic modeling of mammalian cell culture bioprocessing: the quest to advance biomanufacturing. Biotechnol J 2018;13(3)e1700229. http://dx.doi.org/10.1002/biot.201700229.
[65] Chen C, Le H, Follstad B, Goudar CT. A comparative transcriptomics workflow for analyzing microarray data from CHO cell cultures. Biotechnol J 2018;13(3)e1700228. http://dx.doi.org/10.1002/biot.201700228.
[66] Orellana CA, Marcellin E, Palfreyman RW, Munro TP, Gray PP, Nielsen LK. RNA-seq highlights high clonal variation in monoclonal antibody

producing CHO cells. Biotechnol J 2018;13(3)e1700231. http://dx.doi.org/10.1002/biot.201700231.

[67] Maldonado-Agurto R, Dickson AJ. Multiplexed digital mRNA expression analysis profiles system-wide changes in mRNA abundance and responsiveness of UPR-specific gene expression changes during batch culture of recombinant Chinese hamster ovary cells. Biotechnol J 2018;13(3)e1700429. http://dx.doi.org/10.1002/biot.201700429.

[68] Kol S, Ley D, Wulff T, Decker M, Arnsdorf J, Schoffelen S, et al. Multiplex secretome engineering enhances recombinant protein production and purity. Nat Commun 2020;11(1):1908. http://dx.doi.org/10.1038/s41467-020-15866-w.

[69] Nguyen LN, Baumann M, Dhiman H, Marx N, Schmieder V, Hussein M, et al. Novel promoters derived from Chinese hamster ovary cells via in silico and in vitro analysis. Biotechnol J 2019;14(11)e1900125. http://dx.doi.org/10.1002/biot.201900125.

[70] Brown AJ, Gibson S, Hatton D, James DC. Transcriptome-based identification of the optimal reference CHO genes for normalisation of qPCR data. Biotechnol J 2018;13(1). http://dx.doi.org/10.1002/biot.201700259.

[71] Stach CS, McCann MG, O'Brien CM, Le TS, Somia N, Chen X, et al. Model-driven engineering of N-linked glycosylation in Chinese hamster ovary cells. ACS Synth Biol 2019;8(11):2524–35. http://dx.doi.org/10.1021/acssynbio.9b00215.

[72] Kawabe Y, Komatsu S, Komatsu S, Murakami M, Ito A, Sakuma T, et al. Targeted knock-in of an scFv-Fc antibody gene into the hprt locus of Chinese hamster ovary cells using CRISPR/Cas9 and CRIS-PITCh systems. J Biosci Bioeng 2018;125(5):599–605. http://dx.doi.org/10.1016/j.jbiosc.2017.12.003.

[73] Zhao M, Wang J, Luo M, Zhao M, Han L, Zhang M, et al. Rapid development of stable transgene CHO cell lines by CRISPR/Cas9-mediated site-specific integration into C12orf35. Appl Microbiol Biotechnol 2018;102(14):6105–17. http://dx.doi.org/10.1007/s00253-018-9021-6.

[74] Huhn SC, Ou Y, Kumar A, Liu R, Du Z. High throughput, efficacious gene editing & genome surveillance in Chinese hamster ovary cells. PLoS One 2019;14(12):e0218653. http://dx.doi.org/10.1371/journal.pone.0218653.