Design of Optimal CHO Protein N-glycosylation Profiles

Amann, Thomas

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Design of Optimal CHO Protein N-glycosylation Profiles

Philosophiae Doctor (Ph.D.)
Thesis

Thomas Amann

The Novo Nordisk Foundation Center for Biosustainability

The Technical University of Denmark

August 2018

Supervisor: Professor MSO Mikael Rørdam Andersen
Co-Supervisor: Co-PI Helene Faustrup Kildegaard

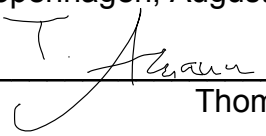
Preface and acknowledgements

My PhD started on September 1st 2015 and was carried out at the Novo Nordisk Foundation Center for Biosustainability (CfB) within the Technical University of Denmark. The project was hosted by the Section for CHO Cell Line Engineering and Design and financed by the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No. 642663 and the Novo Nordisk Foundation (NNF10CC1016517). Between January and May 2018 I had my industrial secondment at AGC Biologics.

During the three years of my PhD I always felt grateful to carry out my research in such an inspiring environment where I could develop myself personally and also professionally. The collaborative, talented and friendly colleges at CfB all contributed to this once-in-a-lifetime experience.

Starting three years ago it became very fast visible, there will not only be ups but also downs. The emotional rollercoaster was intense but is a well known part of science as I know now. However, the good times exceeded the low times by far. This was also due to the outstanding guidance of my supervisors Helene and Mikael, I am very thankful towards you two, further for your trust in me and the freedom you gave me during the different projects.
I want to point out that I am very grateful to have been part of the eCHO program, and I thank everyone who participated therein, it was such a pleasure to get to know so many fellow young researchers with different backgrounds and personalities. Special thanks go to Ankita, Sara, Nuša, Julie and Lise who I had the most productive discussions with and who were always there to celebrate successful times but also to support me during challenging periods. I thank Henning, Daniel, Jae Seong, Eric, Gyun Min, Daria, Kai and TK for inspiring me during my projects and being supportive whenever I had problematic scientific questions. Anders, thank you for introducing me to the world of N-glycans and all the successful projects we have collaborated in. Stef, thank you for all the support on the downstream parts of my projects.
I am in particular grateful for the endless help and assistance by Nachon, KK, Zulfiya, Saranya, Sara, Johnny, Karoline, Tune and Helle. Without you my PhD would have been significantly less fruitful and also less fun, thank you so much. I also want to thank the colleges at AGC Biologics for giving me the opportunity to gain industrial experience and for all the interesting discussions we had.
Finally I want to thank my family and friends who were always there when I needed them the most and who tolerated times when I was less available due to the busy years of the PhD. Without you this work would not have been possible.

Copenhagen, August 30th, 2018



Thomas Amann

Abstract

The amount of marketed therapeutic glycoproteins is increasing steadily and so does the knowledge about the importance and effects of N-glycosylation for patient safety, drug efficacy and pharmacokinetics. Unlike many other expression platforms, chinese hamster ovary (CHO) cell lines possess the ability to provide biopharmaceuticals with N-glycans similar to humans and are therefore the preferred expression host for the majority of glycoproteins.

However, the N-glycan profile of CHO is very heterogeneous and only human-similar, but not human-identical. For some therapeutic products a more homogeneous sugar profile with certain human-identical N-glycan structures is desired. Therefore, it is a fundamental aim to re-design the N-glycan machinery of CHO to produce tailored homogeneous N-glycan structures.

The overall purpose of the thesis was to engineer CHO cells towards specific N-glycan structures. To save time during the cell line development, we examined CRISPR/Cas9 multiplexing to target several genes simultaneously. By this we aimed to provide cell lines for the production of biopharmaceuticals with homogeneous product quality and human-identical N-glycan structures.

The first part of the thesis is a review introducing to the topic and displays how genetic engineering tools as CRISPR are widely used for N-glycan engineering in CHO but also other expression platforms. Following this is an explorative study of CRISPR/Cas9 multiplexing ten gene targets in CHO presenting observed advantages and limitations of the applied protocol.

In the main part of the thesis, the successful production of non-galactosylated glycoproteins (mAb and EPO) after generating cell lines with disruption of B4GALT1, 2, 3 and 4 are depicted. The decrease of galactosylation and heterogeneity of N-glycans was also found on total secreted proteins of the developed cell lines. Furthermore, the generation of a cell line with ten gene disruptions and overexpression of a human glycosyltransferase allowed the production of recombinant A1AT and C1INH with human-like N-glycosylation and *in vitro* activity. The generated cell lines allow the study of possibly novel applications for non-galactosylated glycoproteins and a sustainable and safe production platform to provide recombinant A1AT and C1INH. Finally, we conclude and discuss future perspectives of the obtained results in this thesis.

Dansk sammenfatning

Antallet af terapeutiske glykoproteiner tilgængeligt på markedet er støt stigende og ligeledes gør viden om effekten og betydningen af N-glykosylering for patientsikkerhed, lægemiddel-effekt og farmakokinetik. Kinesiske hamsterovarie-celler (CHO) besidder, ulig mange andre ekspressions-platforme, evnen til at forsyne biofarmaceutiske lægemidler/proteiner med N-glykaner der ligner humane og er derfor den foretrukne ekspression svært for størstedelen af glykoproteiner.

CHOs N-glykan profil er yderst heterogen og human-lignende, men ikke human-identisk. Ved nogle terapeutiske proteiner er en mere homogen sukker profil, med bestemte human-identiske N-glykan strukturer, tilstræbt. Derfor er det et fundamentalt mål at omdesigne CHO's N-glykan maskineri for at producere skræddersyede homogene N-glykan strukturer.

Det overordnede formål med afhandlingen var at konstruere CHO-celler med specifikke N-glykan strukturer. Vi undersøgte CRISPR/Cas9-multiplexing for at målrette flere gener samtidigt. Herved tilstræbte vi at tilvejebringe cellelinier til produktion af glykoproteiner med homogen produktkvalitet og human-identiske N-glykan strukturer.

Den indledende del af denne afhandling består af et review, der introducerer emnet om og belyser hvorledes genteknologiske redskaber som CRISPR er vidt anvendt for N-glykan optimering af CHO cellelinjer, men også andre ekspressions-platforme. Efterfulgt af denne, ses et eksplorativt studie om CRISPR/Cas9 hvor målet var simultant at afbryde ti gen-targets i en CHO cellelinie, hvori der også præsenteres observerede fordele og ulemper ved anvendte protokol.

I hovedteksten af denne afhandling afbildes den succesfulde produktion af ikke-galaktosylerede glykoproteiner (mAb og EPO) ved generering af cellelinier med afbrydelse af generne B4GALT1, -2, -3 og -4. Den aftagende N-glykan galaktosylering og heterogenitet blev også afspejlet i de totale secernerede proteiner af de udviklede cellelinier. Derudover tillod generationen af en cellelinie, med ti gen-afbrydelser og overekspression af en human glykosyltransferase, produktionen af rekombinant A1AT og C1INH med human-identiske N-glykosylering og *in vitro* aktivitet. Afslutningsvist konkluderes og diskuteres fremtidige perspektiver af de opnåede resultater i denne afhandling.

List of publications

This thesis is compiled of the following articles and manuscripts (published and unpublished):

I. **Genetic engineering approaches to improve biopharmaceutical quality attributes in different production platforms**
Amann, T. , Schmieder, V. , Kildegaard, H. F. , Borth, N. and Andersen, M. R. (2018). (*Manuscript in preparation*)

II. **Deca CHO KO: exploring the limitations of CRISPR/Cas9 multiplexing in CHO cells**
Amann, T. , Hansen, A. H. , Lee, G. M. , Andersen, M. R. and Kildegaard, H. F. (2017). (*unpublished work*)

III. **CRISPR/Cas9-Multiplexed Editing of Chinese Hamster Ovary B4Gal-T1, 2, 3, and 4 Tailors N-Glycan Profiles of Therapeutics and Secreted Host Cell Proteins**
Amann, T. , Hansen, A. H., Kol, S. , Lee, G. M., Andersen, M. R. and Kildegaard, H. F. (2018). *Biotechnology Journal*. doi:10.1002/biot.201800111

IV. **Glyco-engineered CHO cell lines producing alpha-1-antitrypsin and C1 esterase inhibitor with fully humanized N-glycosylation profiles**
Amann, T. , Hansen, A. H. , Kol, S. , Hansen, G. H. , Arnsdorf, J. , Nallapareddy, S. , Voldborg, B. , Lee, G. M. , Andersen, M. R. , Kildegaard, H. F. (2018). (*submitted manuscript*)

Table of contents

Thesis objectives and structure

A steady increase in the demand of recombinant biopharmaceuticals for the treatment of numerous diseases has been observed during the last decades[1]. To meet these demands there are several cell platforms available, which include yeast cells, insect cells, plant cells as well as mammalian cells and bacterial expression systems. Many recombinant biopharmaceuticals require post-translational modifications to fulfill their *in vivo* activities or to prevent immunological reactions or fast clearance in the human body. N-glycosylation is undisputed the single most important post-translational modification and has a sincere influence on immunogenicity, pharmacodynamics and -kinetics of biopharmaceuticals[2]. The choice of expression system has a direct impact on N-glycosylation and varies tremendously between plants, bacteria, yeast and mammalian expression systems. Mammalian expression systems display a human-like glycosylation machinery. Recombinant proteins produced in mammalian cell systems are therefore similarly glycosylated as native human proteins. Chinese hamster ovary (CHO) cells are the predominant working horse for the production of biopharmaceuticals within mammalian expression platforms[1]. Advantages of working with CHO cells are relatively easy scale-up, efficient DNA transfection, long history of approved biopharmaceuticals and N-glycosylation close to humans. However CHO N-glycosylation harbours two main challenges, **(i) heterogeneous N-glycosylation** leading to inconsistent product quality and **(ii) missing human-like sialylation**, both being addressed in this thesis.

Attempts of controlling heterogeneous CHO N-glycosylation are often made by process or medium design[3,4]. These attempts are however limited and can not generate biopharmaceuticals with a single predominant N-glycan structure. The availability of genetic engineering tools allows the design of cells with specific N-glycan structures by disrupting the DNA sequence of undesired glycosyltransferases, the enzymes of the N-glycan machinery. With the usage of zinc-finger nucleases (ZFN), Yang et al. successfully demonstrated a protocol to generate CHO cells with designed homogeneous N-glycan structures, mostly by step-wise disruption of several glycosyltransferases[5].

However, stacking gene knockouts in several steps is time-consuming and laborious leading to **(iii) long and complex CHO cell line development (CLD) protocols**, another challenge when working with CHO cell line engineering.

The presence of glycosyltransferase-isoforms is furthermore contributing to the need of stacking several targets to achieve homogeneous N-glycans via target gene disruption. Moreover, those isoforms represent an additional challenge for the design of homogeneous N-glycans in CHO. The **(iv) level of contribution of each isoform** to the overall CHO N-glycosylation of proteins is not completely understood yet. It is furthermore necessary to investigate if designing certain N-glycan structures via the disruption of glycosyltransferases **(v) interferes with CHO cell growth.** Thereby the challenges of CHO N-glycan engineering are listed as:

(i) presence of heterogeneous N-glycosylation
(ii) lack of human alpha-2,6-sialylation
(iii) long and complex CLD protocols
(iv) presence of glycosyltransferase-isoforms
(v) effect of N-glycan engineering on cell growth

The thesis approaches above mentioned challenges by identifying suitable gene targets and by developing protocols for efficient gene disruption via CRISPR/Cas9 multiplexing. It has previously been shown that up to three CHO genes can be disrupted simultaneously applying CRISPR/Cas9[6]. Such a multiplexing approach allows a faster disruption of several target genes compared to repeated rounds of gene target disruptions via ZFN technology as described earlier. By CRISPR/Cas9 multiplexing it is furthermore possible to investigate the contribution of CHO glycosyltransferase-isoforms on overall N-glycosylation.

Figure 1 outlines how above challenges are approached by the objectives of the thesis, which are:

1) **Providing an overview of gene targets for N-glycan engineering in CHO cells**
2) **Shortening CHO CLD by developing CRISPR/Cas9-multiplexing protocols for simultaneous KO of >3 gene targets**
3) **Multiplexing glycosyltransferase-isoforms to reduce N-glycan heterogeneity, study their contribution to N-glycosylation and analysis of cell growth**
4) **Mimicry of human glycosylation on recombinant human glycoproteins expressed in CHO cells and analysis of cell growth**

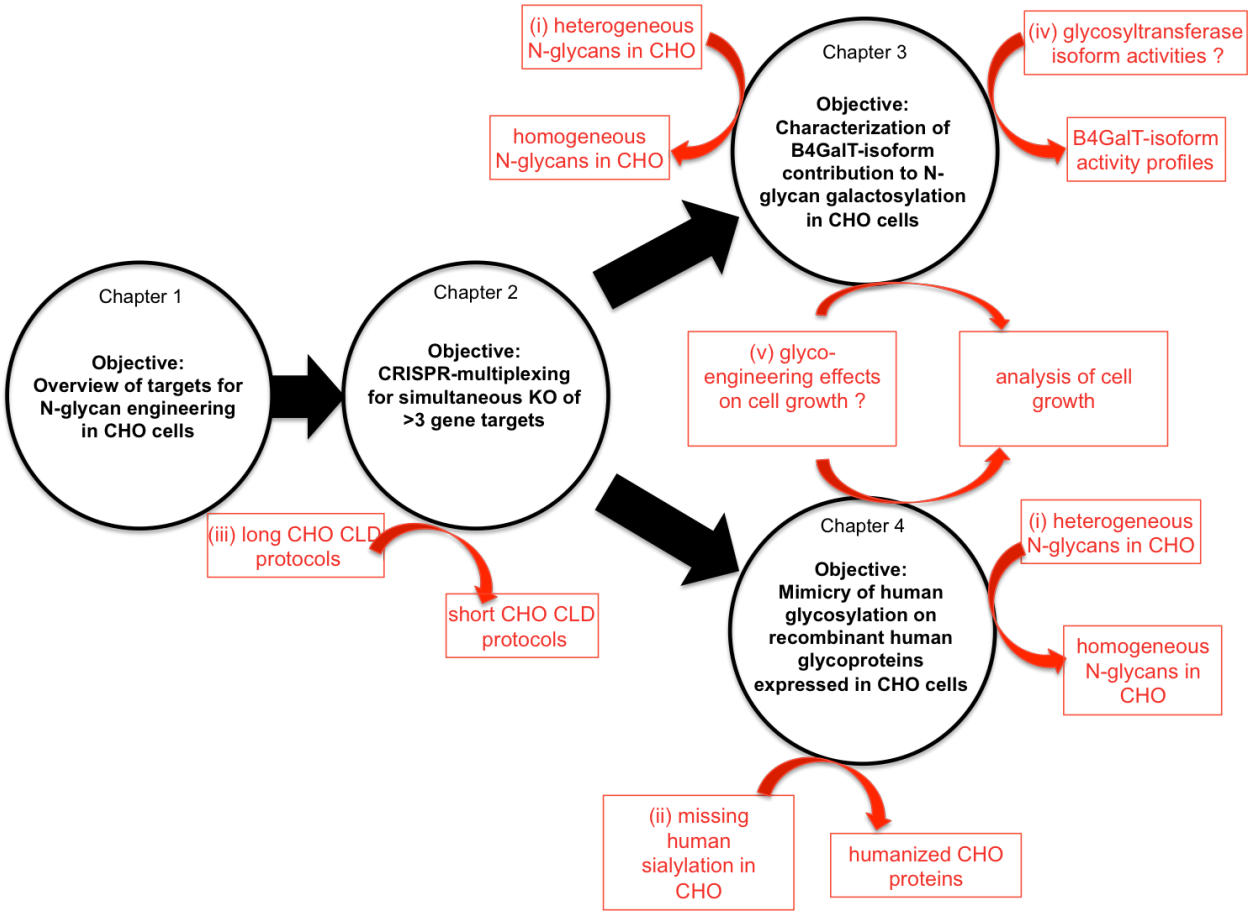


Figure 1: Objectives of the thesis. Flowchart of the thesis objectives within the thesis chapters 1 - 4 (round circles). The approached challenges (i) - (v) are illustrated in red and partially covered by more than one chapter.

The four parts of the thesis all focus on the overarching goal of my project: design and development of CHO cell lines with optimal N-glycan profiles. **Chapter 1** is a review introducing to the field of genetic engineering in recombinant production platforms and discusses engineering approaches towards improved N-glycosylation and other product quality attributes. It includes an overview of suitable gene targets for N-glycan engineering and other PTM. The target list became a valuable source for the selection of gene targets within the work of Chapters 2 - 4. The study presented in **Chapter 2** was an explorative attempt to investigate the feasibility of knocking out ten targets by simultaneous CRISPR/Cas9 multiplexing. The objective of this chapter was mainly to reduce the length of CHO CLD by providing a protocol for the simultaneous disruption of more than three gene targets. We expected to gain insight into the advantages and drawbacks of targeting a high number of genes. The work of **Chapter 3** was based on the learnings from Chapter 2. We successfully disrupted up to four galactosyltransferase-isoforms by applying the protocol from Chapter 2. The objective was to investigate the enzymatic activities of the targeted isoforms. Thereby we also aimed to produce therapeutic proteins with decreased N-glycan galactosylation and N-glycan heterogeneity. Examination if decreased galactosylation and decreased N-glycan heterogeneity is interfering with CHO cell growth was a further goal of this chapter. In **Chapter 4** the objective was to design CHO cell lines capable of producing recombinant proteins with a fully humanized and homogeneous N-glycan profile. Based on the target list of Chapter 1 and the multiplexing protocol of Chapter 2 we aimed to generate a 10x knockout cell line producing recombinant human plasma proteins with fully humanized N-glycan structures and *in vitro* activities similar to the human counterparts. Fully humanized alpha-2,6-sialylation was designed by disrupting CHO-specific sialyltransferases and stable integration of human ST6 beta-galactoside alpha-2,6-sialyltransferase 1. Identification of clones with correct N-glycosylation by lectin staining was a further part of the design. The final **Chapter 5** contains concluding remarks about the results and future perspectives for the research presented in this thesis. The research and results presented in this thesis were completed in a collaborative effort between me and my colleagues.

1. Walsh, G. Biopharmaceutical benchmarks 2014. *Nat. Biotechnol.* **32,** 992–1000 (2014).
2. Liu, L. Antibody glycosylation and its impact on the pharmacokinetics and pharmacodynamics of monoclonal antibodies and Fc-fusion proteins. *J. Pharm. Sci.* **104,** 1866–1884 (2015).
3. Fan, Y. *et al.* Amino acid and glucose metabolism in fed-batch CHO cell culture affects antibody production and glycosylation. *Biotechnol. Bioeng.* **112,** 521–535 (2015).
4. Spearman, M., Rodriguez, J., Huzel, N., Sunley, K. & Butler, M. Effect of Culture Conditions on Glycosylation of Recombinant beta-Interferon in CHO Cells. in *Cell Technology for Cell Products* 71–85
5. Yang, Z. *et al.* Engineered CHO cells for production of diverse, homogeneous glycoproteins. *Nat. Biotechnol.* **33,** 842–844 (2015).
6. Grav, L. M. *et al.* One-step generation of triple knockout CHO cell lines using CRISPR/Cas9 and fluorescent enrichment. *Biotechnol. J.* **10,** 1446–1456 (2015).

CHAPTER 1

Genetic engineering to improve the quality of biopharmaceuticals

This chapter introduces different genetic engineering tools, recombinant production platforms and discusses engineering approaches to improve important protein quality attributes whereof N-glycosylation is the main subject. We highlight the importance of CRISPR/Cas9 in the field of cell line development and how the technology can boost N-glycan engineering to aim for cell lines with designed homogeneous N-glycan profiles which is the overall objective of this thesis.

Genetic engineering approaches to improve biopharmaceutical quality attributes in different production platforms

Thomas Amann[1], Valerie Schmieder[2,3], Helene Faustrup Kildegaard[1], Nicole Borth[3], Mikael Rørdam Andersen[4]

[1]The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark
[2]ACIB GmbH, Austrian Center of Industrial Biotechnology, Graz, Austria
[3]Department of Biotechnology, BOKU University of Natural Resources and Life Sciences, Vienna, Austria
[4]Department of Biotechnology and Biomedicine, Technical University of Denmark, Kgs. Lyngby, Denmark

---

Abstract

The number of approved biopharmaceuticals is increasing steadily while product quality attributes remain of major importance. Within the available variety of expression hosts, the production of biopharmaceuticals faces diverse limitations for post-translational modifications (PTM). However, different biopharmaceuticals demand a different degree of certain PTMs for proper functionality. Besides PTMs, product quality can also be affected by host-specific proteases, product aggregation or undesired impurities. With the growing toolbox of genetic engineering technologies it is possible to address general, host- or biopharmaceutical-specific product quality obstacles.
In this review we present diverse expression systems derived from mammalian, bacteria, yeast, plants, and insects as well as available genetic engineering tools. We list genes for knockout / knockdown and overexpression from meaningful approaches improving biopharmaceutical quality attributes and discuss their applicability as well as future trends in the field.

Keywords: Chinese Hamster Ovary, Cell Engineering, PTM, Expression system, CRISPR

1. Introduction

The demand for recombinant biopharmaceuticals is steadily increasing and so does the proportion of molecules which harbor post-translational modifications (PTMs) e.g. monoclonal antibodies (mAbs) and blood factors[1]. With increasing size of the desired product, mostly also the amount of PTMs is increasing too, giving rise to highly complex proteins. Correct PTMs, in particular N-glycosylation, can be substantial for *in-vivo* functions, appropriate plasma half-lifes or prevent immunological reactions after injection. While for instance non-glycosylated drugs are mostly produced in bacteria and yeast[2], glycosylated proteins demand expression in mammalian cell lines. Therefore, the increasing number of approved biopharmaceuticals produced in mammalian expression systems aligns with the trend for the growing need for proper PTMs and other quality attributes. However, microbial production is quantitatively still dominating and both expression systems, mammalian and nonmammalian, represent certain advantages but also challenges as reviewed previously[2,3]. At the same time plant-based and insect cell expression systems are catching up slowly.

Many cell platforms are engineered towards human-like glycosylation since human N-glycan profiles are preferred on many therapeutics[4]. At the same time certain therapeutic proteins with for humans untypical N-glycosylation are beneficial too[5]. Mimicking other human-like PTMs as gamma (γ)-carboxylation, phosphorylation and tyrosine sulfation can also be challenging during the production of therapeutic proteins. Several regulatory agencies across the world like the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA) consider quality attributes, especially glycosylation, as one of the most critical requirement during biopharmaceutical production[6-11]. In addition, product degradation and undesired impurities often hinder high-yield processes during cell cultivation or downstream protocols.

The increase of characterized and sequenced host cell platforms as well as efficient genetic engineering tools are speeding up the process to produce protein products with elevated qualities. In this review we present (i) the variety of expression platforms with their strengths and drawbacks, (ii) available genetic engineering tools and (iii) a comprehensive retrospect of successful synthetic biology approaches to enhance product qualities.

2. Available expression platforms for the production of different biopharmaceuticals

Nowadays, numerous pro- and eukaryotic cells are available for the expression of biopharmaceuticals in industrial scale. Those expression platforms can be mainly grouped into mammalian, bacteria, yeast, plant and insect systems[12,13]. In addition to correct post-translational modification of the desired bioproduct, the selected species and production hosts have to fulfil requirements to meet the needs of bioprocessing. In detail, accessibility towards genetic engineering with an available molecular tool box, genomic stability, a functional secretion machinery, high growth and productivity rates, easy bioprocess handling, downstream purification, and scalability are beneficial features of production systems[14]. Additionally, the history of biopharmaceutical approvals produced in a certain cell line plays an important role when it comes to the process of choosing the right expression platform[1,15]. In the following section, different cellular platforms are described with their advantages and disadvantages, highlighting why they are used for the manufacturing of certain biopharmaceuticals.

2.1 Mammalian expression platforms

The class of mammalian expression platforms is divided into human and non-human derived cell lines with Chinese Hamster Ovary (CHO) being the most dominant manufacturing system[16]. Human Embryonic Kidney 293 (HEK293) and CEVEC's Amniocyte Production (CAP) cells are examples of cell lines isolated from humans[17,18]. The main advantage of human cell lines is the potential for human identical PTMs[18]. Non-human cell lines commonly used are for instance CHO, Baby Hamster Kidney (BHK), and murine myeloma cells like NS0, which all have the ability to perform human-like PTMs[15]. Furthermore, these and in particular CHO have the benefit of being resistant towards human viruses resulting in reduced viral inactivation and clearance efforts during downstream processing[19]. Since sufficient protein modulation, folding and secretion can be achieved in mammalian cells, they are the prefered choice for the production of large and complex biopharmaceuticals[15]. Mammalian cells are mostly growing adherently in their native form with moderate growth rates and requiring complex media in some cases supplemented with Fetal Calf Serum (FCS)[20-22]. Indeed, this generally increases the cost as well as the complexity of bioprocessing. Those limitations are overcome by the use of CHO cells, which can be adapted to serum-free suspension growth with a doubling rate of 20-24 h depending on the process[23,24]. Also the long list of approved therapeutics manufactured in CHO and processes that were established for this cell line during the last decades are emphasizing why CHO is the working horse of biopharma[1]. Nevertheless, genetic instability and intensive cell line generation as well as process development timelines are the major bottleneck of this system[25-27].

2.2 Bacterial expression systems

Bacterial expression systems, mainly *Escherichia coli* (*E. coli*), are widely used for the generation of smaller drugs including hormones, growth factors and even more innovative therapeutic agents like peptibodies / peptide-fragment crystallizable region (Fc) fusions[28,29]. Unlike mammalian cells, bacterial microorganisms miss the ability to perform most of the PTMs, including glycosylation, predominantly essential for the production of active biopharmaceuticals[30,31]. Additionally, correct protein folding as well as proper secretion is not a given in most bacterial species leading to the formation of inclusion bodies[32]. Using an appropriate signal sequence, the product of interest can be directed to the periplasm where a reducing environment allows protein oxidation for disulfide bond formation, thus, enabling folding[33,29]. However, this approach is only realizable in gram-negative bacteria[33]. Notably, *Bacillus subtilis* possess very efficient secretion routes due to the lack of the outer cell membrane as reviewed previously[34]. In short, high expression levels, less expensive and less complex bioprocessing, as well as well characterized genetics giving rise to the establishment of a extensive molecular tool box are the benefits of manufacturing in bacteria[35].

2.3 Yeast expression platforms

Pichia pastoris (*P. pastoris*) and *Saccharomyces cerevisiae* (*S. cerevisiae*) are the most important representatives of the yeast protein expression platforms. Yeast microorganisms have the ability to perform typical eukaryotic PTMs including N- and, to some extent, O-linked glycosylation, ubiquitination, sumoylation, and myristoylation[36,37]. Heterologous protein secretion can be achieved by the use of an appropriate signal peptide[38,39]. In contrast to N-glycosylation in mammalian cells, yeast perform mammalian-atypical hyper-mannosylation, and lack the ability of generating sialylated N-glycans[40]. Nevertheless, glycosylation of the biopharmaceutical facilitates further protein processing and secretion[41]. Similar to bacterial expression systems, well characterized genomes, the availability of a molecular and synthetic tool box, moderately rapid and cost-efficient expression, and the absence of endotoxins as well as human virus-resistance are emphasizing yeast as a host for biologics manufacturing of mainly smaller proteins, hormones and vaccines[41-43].

2.4 Plant-derived expression systems

Plant-based systems are in the focus for adaptation as alternative hosts for the production of biopharmaceuticals for many years where the trend has moved to rather use plant cells than the whole plant or plant tissues/organs[44,45]. Therefore, bioprocessing in plant cell lines e.g. tobacco cultivars Bright Yellow 2 (BY-2), *Nicotiana tabacum* 1 cell (NT-1), carrot or rice cells combines the advantages

from plant farming and mammalian manufacturing procedure[46]. On one hand, advantages of the system are growth in simple medium, good scalability, and the given safety since plant cells are not affected by human pathogens. On the other hand, complex downstream purification and low product yields remain to be the major bottlenecks of this host class[46-47]. Plant platforms are engineered to express complex Abs, low molecular weight products like enzymes, hormone, growth factors and cytokines or macromolecules like virus-like particles (VLPs)[45,48]. Hence, transformation of the gene of interest (GOI) and/or engineering tools is done by Agrobacterium infiltration[49]. So far, only one biopharmaceutical recombinantly produced in carrot-derived cells got approval by FDA[45,50]. In plants the performance of PTMs, correct assembly and folding of complex therapeutics, as well as their secretion into the cell culture medium is possible[47]. In terms of N-glycosylation, biopharmaceuticals generated from plant-based systems show plant-specific glycan structures and lack mammalian-specific sialyl acids, which can result in immunogenicity and poor pharmacokinetics, respectively[47,51,52]. Another drawback of manufacturing in plant hosts is the relatively high level of bioproduct proteolysis induced by endogenous proteolytic enzymes, indeed, lowering recombinant protein stability and yields[52-54].

2.5 Insect expression platforms

For the production of recombinant biopharmaceuticals, insect cells go hand in hand with baculovirus infection building the baculovirus-insect cell system (BICS)[55,56]. *Trichoplusia ni* from cabbage looper or the “gold standard” *Spodoptera frugiperda* (Sf), derived from fall armyworm, are examples for insect cell lines[55]. Since the BICS is kind of a “plug-and-play” system in which only the baculovirus vector has to be modified for each product prior transduction of the insect cells, BICS is flexible, fast, and simple to use[55,56]. Even the simultaneous production of several protein subunits is realisable[57]. Moderate expression and difficulties during scale-up are reasons why production in insect cells is still quite rare, thus improving cell line development and bioprocessing are still ongoing in this field[58]. Mainly vaccines / VLPs are generated from insect cell lines, with Cervarix, a vaccine against different cancer types caused by human papillomavirus, as the first drug being approved in 2009[59]. Especially for vaccine manufacturing the BICS is of benefit because the turnover to produce and approve new vaccines is streamlined, thereby assessing the annually need of vaccines due to the seasonally and epidemic character of viruses[60]. As for most eukaryotic expression platforms, insect cells are capable to provide different PTM, whereas protein processing, folding, and secretion of secreted and membrane-bound products are limited and need to be initiated[54,61]. Therapeutics often occur insoluble and poorly processed in the cytoplasm of the insect cells[61]. Similar to yeast expression systems, insect cells build

simple, inconsistent paucimannose N-glycan structures[58]. The low expression of most glycosyltransferases in combination with the presence of many active glycosidases leads to the incompatibility of generating human-like N-glycans in non-engineered insect cell lines[62]. Furthermore, reduced cell viability and declined titers are often an effect from proteolysis induced by baculovirus-derived enzymes, especially in later phases of cultivation[63]. Recently, a insect-specific virus contamination problem caused by the rhabdovirus was reported. Although no evidence that the contamination might be harmful for humans was communicated so far, the biosafety of the BICS has to be questioned and investigated[55].

3. Reported complications within critical product quality attributes

The number of approved biopharmaceuticals is increasing steadily. The predominant indications for biopharmaceuticals are cancer, inflammation-related conditions, hemophilia, metabolic disorders, diabetes and vaccinations against various infectious diseases[1]. The classes of recombinant biopharmaceuticals include monoclonal antibodies (mAbs), hormones, blood-related proteins, enzymes, vaccines, fusion proteins, growth factors and interferons where the distribution of necessary PTMs for proper pharmacodynamics and -kinetics varies within and between the classes.

3.1 Product glycosylation

Recombinant hormones, growth factors and interferons as insulin, recombinant human (rh) growth hormone (rhGH), rh granulocyte-colony stimulating factor (rhG-CSF) and interferon-α (IFN-α) demand rather uncomplex PTMs and therefore can be produced in *E. coli* or *S. cerevisiae*[1,64]. An example is the *E. coli*-derived IFN-α with missing O-glycosylation, however, the bioactivity is comparable to the human counterpart[65]. Similarly, most recombinant vaccines do not need human-like PTMs to fulfill their functions and are predominantly produced in *S. cerevisiae*[1].

Recombinant DNAse (enzyme), rh C1 inhibitor (rhC1INH) (plasma protein), etanercept (Immunoglobulin G (IgG)-Fc fusion protein) are examples for proteins from different product classes which require very specific and complex human-like N-glycans and are therefore produced in mammalian expression platforms[4,66-68].

Within mammalian hosts, the choice of cell platform can also directly impact immunogenic properties of the therapeutic protein. As recently reported, rh coagulation factor (rhF) VIII expressed in BHK cells is more immunogenic than that produced in CHO cells[69].

Rh erythropoietin (rhEPO) demands a high degree of sialylation and branching to achieve the desired biological potency[70]. As for rhFVIII, CHO platforms are the preferred host to achieve human-like N-glycosylation on rhEPO[1].

Correct glycosylation is also important in the main product class of biopharmaceuticals, mAbs. The composition of the sugar moieties often has a direct impact on pharmacokinetics, -dynamics and immunogenicity[71] and is therefore a major target for cell engineering.
An example for a product that does not need complex N-glycosylation, but instead requires mannose N-glycans is the recombinant enzyme glucocerebrosidase. Produced in mammalian cell lines, N-glycans of purified glucocerebrosidase are too complex and need downstream enzymatic treatment to ensure the biological uptake. Expression of rh glucocerebrosidase in engineered carrot root cell culture makes this *in vitro* treatment unnecessary[72].

3.2 Carboxylation, phosphorylation and tyrosine sulfation

However, meeting certain N-glycosylation profiles is not the only PTM challenge faced by many biopharmaceutical producers. For some proteins of the class of blood related products e.g. rhFIX and Protein C, γ-glutamyl carboxylation is crucial for efficacy and proper pharmacokinetics and represents a challenge during recombinant production[73-76].
Compared to N-glycosylation, phosphorylation and tyrosine sulfation are rather unexplored PTMs although phosphorylation is very common and more than 2100 mammalian proteins are predicted to harbor sulfotyrosine[77]. Since natural *E. coli* and other prokaryotic expression platforms are not capable to perform tyrosine sulfation, mammalian cell systems are chosen to produce biopharmaceuticals with needed sulfotyrosine as the case for rhFVIII[78].

3.3 Product aggregation and degradation

Although protein expression in prokaryotic systems as *E. coli* has advantages for some products it can also lead to undesired aggregation of recombinant proteins. The incompatibility of folding proteins of eukaryotic origin and the lack of compartmentalization, PTMs and suitable chaperones are the major contributors to protein aggregation[79]. In general, protein aggregation is less a problem in mammalian expression platforms. However, protein products from mammalian cells can face proteolytic degradation before secretion or due to the presence of host proteases in the cell culture medium[81-84]. This in particular is a major hurdle for the development of an increasing number of non-antibody formats but can be addressed by host cells with protease gene deletions[85].

To address the hurdles for desired product quality, it is important to choose a suitable type of expression host which is capable to perform needed protein modifications without compromising product yield or other economical factors. Lacking capabilities however can be encountered by several available genetic engineering tools which are described in the next section.

4. Genetic engineering tools

For reasonable and sufficient genetic engineering of cells, knowing the genomic sequence as well as the expression level of endogenous proteins and their functions is of fundamental importance[86-91]. Therefore, the rapid improvements in terms of next-generation sequencing (NGS) techniques[92] and data generation via several omics approaches including all cellular levels from genome, epigenome, transcriptome, proteome and reactom, are of benefit when talking about cell line editing for the proper post-translational modulation of desired biopharmaceuticals[93-95].

Host cell lines are engineered by the overexpression (OE) or disruption of individual enzymes or proteins of entire pathways to mediate naturally occurring PTMs to in the end being able to produce functional biopharmaceuticals[96-98]. Following this approach, the GOI are integrated randomly into the genome resulting in heterogeneous cell populations, and most likely flux imbalances in the cells[99]. With numerous omics data and functional studies arising, while at the same time cell line engineering techniques are improving rapidly, endogenous PTM alteration to meet biopharmaceutical-specific requirements becomes more and more feasible in a targeted manner[100-105]. Nowadays, several molecular editing tools are available for this purpose. Treating cells with non-coding RNAs (ncRNAs) like small hairpin RNA (shRNA), microRNA or small interfering RNA (siRNA) is one way of control on transcriptomic level without affecting the genome[106-109]. Those RNA molecules can either bind the mRNA target perfectly, which results in direct mRNA cleavage, or with incomplete complementary conducting translational repression and further mRNA degradation[110]. Thus, both mechanisms induce gene down regulation, respectively. Since engineering by microRNA or RNA interference (RNAi) approaches impacts the post-transcriptional cell stage, the resulting effect is only temporary as long as the RNA molecules are present in the cells. Furthermore, some ncRNAs have multiple targets, which could be of benefit when multiple enzymes need to be altered simultaneously, or also being a major bottleneck of the system if unwanted side-effects occur[111]. Moreover, genome editing methods including meganucleases, zinc finger nucleases (ZFN) and transcription activator-like effector nucleases (TALEN) have emerged in the field. These endonucleases can be programmed to build a DNA-enzyme-complex at any desired spot in the genome inducing a double-stranded-break (DSB), which will be repaired by the cell's own mechanisms[112-114]. DSBs are either corrected by the error-prone non-homologous end joining (NHEJ) pathway resulting in small insertions/deletions (InDels) of nucleotides, thus, most likely in a frame-shift/loss-of-function mutation[115] or in the presence of a suitable homology-mediating template the DNA cut is repaired through the homology directed repair (HDR) pathway[116]. Inducing the NHEJ mechanism is the prefered way if one wants to knockout (KO) a gene and the corresponding function whereas initiating HDR is the method of choice to introduce new

sequences into the host's genome[117,118]. In contrast to NHEJ, sequence knockin (KI) via HDR is occuring at a much lower frequency in most if not all available expression systems[119,120]. Meganucleases, TALENs and ZFNs are protein-based genome engineering techniques. Thereof, for each target site new proteins have to be designed and produced, indeed, making the application of those tools time-consuming, expensive and laborious[121]. Here, the clustered regularly interspaced short palindromic repeats (CRISPR) system has been developed lately as a new technology for targeted genome editing. In nature, the RNA-protein-complex is part of the the adaptive immune system of different bacterial strains to protect these from an infection caused by bacteriophages[122]. The original tool consists out of the CRISPR-associated (Cas) endonuclease - with Cas9 from *Streptococcus pyogenes* as the most extensively used derivative - and a short RNA molecule, which is guiding the CRISPR enzyme to any DNA sequence to specifically induce a molecular change at the 5'-primer region upstream of a protospacer adjacent motif (PAM), an essential component for CRISPR activity and target recognition[123]. The guiding RNA (gRNA) molecule consists of two parts. (I) A CRISPR RNA (crRNA) sequence containing the variable target sequence (20 nt) followed by a repetitive part (nt). (II) A trans-acting CRISPR RNA (tracrRNA) sequence, which is binding the crRNA repeats by their complementarity[124]. By doing so, a hairpin structure is formed leading to the attraction of the CRISPR enzyme[125]. In 2012, Jinek et al. implemented a synthetic whole-in-one transcribed single guide RNA (sgRNA) for CRISPR/Cas9[122]. Since only the guide sequence of the molecule has to be modified for each new target gene, the technology is very flexible, easy to handle, and the application forms are diverse[126]. By the synchronous delivery of a sgRNA pair, entire gene bodies or even non-coding regions can be removed from the cell system allowing proper gene function characterization without a potential interference of truncated protein versions and the study of the untranslated part of the genome[127-129]. So far, CRISPR/Cas9 has successfully been used in a wide range of organisms and cell lines for genome alteration by stable gene KO or KI. Additionally, reversible gene activation as well as repression can be achieved by using a catalytically dead Cas9 (dCas9) fused to an effector domain[130]. Currently, this technique enables for instance gene transcription activation or deactivation by de- or methylation of CpG islands within a promoter[131-133]. Moreover, during the last years the Cas9 alternative enzyme CRISPR-associated endonuclease in *Prevotella* and *Francisella* (Cpf1) - but mainly from the two organisms *Acidaminococcus sp BV3L6* and *Lachnospiraceae bacterium ND2006* - revealed its high potential for further accelerating the CRISPR tool box[127-134]. Since Cpf1 recognizes a PAM sequence (5'-TTTN-gRNA) different than Cas9 (5'-gRNA-NGG), one can switch between the two CRISPR systems depending on whether the sequence to be tackled is T- or G-rich. Also, Cpf1 requires a shorter sgRNA and has an additional RNase activity allowing to transcribe sgRNAs from an array if simultaneous editing of multiple genes is requested[135]. A current trend to enhance the CRISPR

technology is the combination of various techniques. In 2015 Lee et al. showed a boosted HDR-mediated KI in CHO by additionally using classical cell line development methods[136]. Another innovative CRISPR approach was demonstrated by Eisenhut et al. in 2018 allowing multiple gene activation by the subsequently deletion of repressor elements using sgRNA pairs[137]. CRISPR-based genetic screenings are the state-of-the-art approach to investigate cellular processes or identify new potential cell line engineering targets as reviewed by Shalem et al., in 2015[138]. Therefore, a subset of proteins up to the entire proteome are covered by an appropriate number of gRNAs to create gene disruptions[139,140]. Moreover, by applying gRNA pairs genomic deletions are induced allowing the characterization of the untranslated part of the genome[141]. By using a gRNA library along with dCas9, coding or non-coding genes can be activated or inactivated[142]. After altering the genome, one can select for a desired phenotype and investigate the changes introduced on the genomic level. Nowadays, numerous CRISPR libraries are commercially available, however, online tools are accessible to design and generate customized gRNA libraries, respectively[143]. Genetic screens are performed in array[144] or pooled format[145]. Arrayed CRISPR screens require automatized plate processes and thus are complex to handle. In comparison, genetic screens using CRISPR in pooled format can be easily performed in a controllable high-throughput manner[146]. Overall, CRISPR screens are powerful methods to link the genome to a generated phenotype.

5. Genetic engineering improves product quality attributes

Traditional approaches to improve product quality are based on medium and process design in addition to *in vitro* treatment of the purified therapeutic protein[147,148]. Depending on the medium additives and enzymes for *in vitro* treatments, these strategies are rather expensive, especially in production scales. With the earlier described genetic tools, it is however possible to change the product quality via targeted engineering of the production host[149]. In this section, we describe selected examples of cell line engineering that contributed to improved product quality in eukaryotic and prokaryotic expression platforms. All described engineering examples are summarized in Table 1, divided by the respective PTM.

5.1 N-glycosylation

Glycosylation is the single most important PTM influencing product quality and may differ by sugar chain length, sequence as well as branching sites and branching numbers[150,151]. Glycosylation is not only species- and cell-specific but also affected by culture conditions[152] causing differences between recombinant human glycoproteins and their endogenous counterparts. The N-glycan composition can

have tremendous impact on immunogenicity, pharmacodynamics and -kinetics of biopharmaceuticals[71]. An example is the lacking ability of yeast, plants and insects to produce complex N-glycans with terminal sialylation[52,153,154], which often has negative consequences for pharmacokinetics[155] and therefore is a major reason for the small number of approved biopharmaceuticals from these platforms[1].

5.1.1 CHO, the main representative of mammalian expression hosts

A major strength of CHO cells is their capability to produce biopharmaceuticals with human-like N-glycans. However, there are attempts to engineer CHO towards a fully humanized N-glycosylation profile. A small proportion of CHO N-glycans for instance consists of N-Glycolylneuraminic Acid (Neu5Gc) which is immunogenetic for humans. Knockdown (KD) of Cytidine Monophospho-N-Acetylneuraminic Acid Hydroxylase (CMAH) with antisense RNA (asRNA) reduced the formation of Neu5Gc[156] (Table 1).

Other approaches to transform CHO N-glycosylation into human N-glycosylation include the expression or activation of the ST6 beta (β)-Galactoside alpha (α)-2,6-Sialyltransferase (ST6GAL) 1, which is encoded in the CHO genome without being active[19]. Overexpression of ST6GAL1 cDNA or stable, but reversible, endogenous ST6GAL1-activation via epigenetic editing with CRISPR are possibilities to generate CHO cell lines with humanized sialylation profiles[157,158].

Besides approaches towards humanized N-glycans, it is well known that certain sugar residues can improve the activity and/or serum half-life of particular biopharmaceuticals. A well-known example therefore is the KO/KD of Fucosyltransferase (FUT) 8 to reduce mAb core-fucosylation for significantly enhanced antibody-dependent cell-mediated cytotoxicity (ADCC) which was achieved by applying different genetic engineering tools (including ZFN, CRISPR and siRNA)[159-161]. Similarly, the clearance rate of rhEPO in the human body is affected by the degree of sialylation where under-sialylated EPO is degraded in the liver[162]. Zhang et al. reported enhanced IFN-γ sialylation after sh- and siRNA silencing the genes N-Acetyl-α-Neuraminidase (NEU) 1 and 3, making them promising KO targets to generate a host cell for the production of highly-sialylated EPO[163].

To address heterogeneous product N-glycosylation in mammalian cells, which is particularly a problem in CHO cell lines, different groups applied ZFN-technology and CRISPR/Cas9 for simultaneous multiplexing of several target genes. The researchers successfully engineered CHO cells for homogeneous N-glycans on biopharmaceuticals and additionally investigated the impact of the generated gene disruptions on cell cultivation performance[164,165]. For a detailed insight into CHO N-glycan engineering, we refer to a recently published review[166].

Another successful engineering attempt introduced a plant-derived β1,2-Xylosyltransferase (XYLT) into a CHO cell line to improve vaccine efficacy via N-xylosylation[167] emphasizing that the transfer of certain glycosylation genes between species can result in promising PTM of biopharmaceuticals.

5.1.2 Bacterial expression systems

Many bacterial expression systems including *E. coli* are incapable of protein glycosylation. However, researchers identified bacteria which are able of glycosylation[168-170] and could transfer this ability into the industrially relevant *E. coli*[171] showing for the first time that bacteria can be an alternative source for recombinant glycoproteins. While this first study produced bacterial N-glycans with great structural difference to their eukaryotic counterparts, a more advanced study reported that OE of yeast N-Acetylglucosaminyltransferase (ALG) 1, 2, 13 and 14 resulted in eukaryotic N-glycans on single-chain variable fragment (scFv) from *E. coli*[172]. This system can be used as a starting point to create novel glycoconjugates within bacterial expression platforms.

5.1.3 Yeast expression systems

Yeast high-mannose N-glycans are often engineered towards more human-like N-glycans by first decreasing high-mannose proportions to then introduce genetic elements for complex N-glycosylation. First, hyper-mannosylation is eliminated by disruption of α-1,6-Mannosyltransferase (OCH1), ALG3 and ALG11[173-176] or expressing Mannosidase (MAN) genes[177]. In the second step, glycosyltransferases, -transporters and sugar-synthesis enzymes are introduced (e.g. N-Acetylglucosaminyltransferase (GNT) and II, β-1,4-Galactosyltransferase (B4GALT) 1, ST6GAL, Glucosamine (UDP-N-Acetyl)-2-Epimerase (GNE), Cerebroside Sulfotransferase (CST)). Based on this strategy, Hamilton and co-workers introduced 14 elements and reported yeast strains producing complex glycoproteins with >90% terminal sialylation[178].

Besides engineering towards human-like sugar residues, the reduction of N-glycan macroheterogeneity to ensure homogenous products is also of major interest. OE of Dolichyl-Diphosphooligosaccharide Protein Glycosyltransferase Subunit STT3 (STT3D) has been shown to increase side occupancy and therefore reduced macro-heterogeneity during the production of rh Granulocyte-macrophage colony-stimulating factor (GM-CSF) and mAb[179].

5.1.4 Plant-based expression systems

Plants contain β-hexosaminidases, which give rise to N-glycan structures with terminal mannoses[179,180] and thus are far from their human counterparts. However, this is not only a drawback and can be used

for the production of biopharmaceuticals with efficient drug internalization by mannose receptors in e.g. Gaucher´s disease patients[52]. One approved drug for Gaucher´s disease, rh glucocerebrosidase (GC) (taliglucerase alfa), is produced in carrot root cell culture with mainly xylosylated paucimannose N-glycans[181,182]. Non-engineered plants create the possibly immunogenic sugar residues β1,2-xylose and core α1,3-fucose[180] and can be engineered to lack the two sugar residues via KO/KD strategies. The use of RNAi resulted in the production of rhGC in *Nicotiana benthamiana* (*N. benthamiana)* without immunogenic sugar residues[183] and several mAb formats with engineered N-glycans were produced with the help of siRNA[184], CRISPR/Cas9[185] and RNAi technologies[186].

5.1.5 Insect-based expression systems

Overexpression of human glycosyltransferases MGAT2, B4GALT and ST6GAL1 has been shown to reduce insect-typical oligo- and paucimannose residues and introduced complex type N-glycans on recombinant proteins produced in insect cell lines[187,188]. As for previously described expression systems, CRISPR tools are also used for insect cell lines to design N-glycans more similar to the human counterparts. An example therefore is the disruption of β-N-Acetylglucosaminidase (GalNAc) / *fused lobes* gene (FDL) to produce EPO with reduced oligo- and paucimannose structures in Sf9 cells[189].

5.2 O-glycosylation

Similar to N-glycosylation, O-glycosylation is a very common PTM in mammalian cells, yeast and plants where different types of O-glycosylation (e.g., O-linked fucose, glucose, mannose, xylose, or GalNAc) have been described on secreted proteins[190]. Immunoglobulin A (IgA) class molecules, rhEPO and Etanercept are examples of therapeutic proteins with O-glycosylation. However, the contribution of O-glycan moieties to therapeutic properties is still not well understood and suitable targets for O-glycan engineering are more difficult to identify than for N-glycan engineering as reviewed previously[191]. In mammals, numerous and often differentially expressed glycosyltransferases give rise to complex O-glycan pathways and highly heterogeneous O-glycans[192-194]. In contrast, O-glycan patterns from non-mammalian cell platforms are often more simple and less heterogeneous, allowing for the expression of proteins where non-human like O-glycans are acceptable[195]. On the background of such rather uncomplex O-glycosylation structures in *S. cerevisiae* and plants, researchers successfully built up different core O-glycans by *de novo* engineering as presented in Table 1[196-200]. For some products the O-mannosylation of *P. pastoris* is undesired and can be counteracted by overexpression of an α-mannosidase enabling further elongation with mammalian-type modifications as reported by Hamilton et al[201].

Engineering of O-glycosylation in mammalian cell lines mostly aims to either simplify the sugar moieties to reduce product heterogeneity or to design distinct O-glycans for superior protein functions. Yang et al. reported decreased O-glycan heterogeneity after disruption of Core 1 Synthase Glycoprotein-N-Acetylgalactosamine 3-β-Galactosyltransferase 1 (C1GALT1) Specific Chaperone 1 (COSMC) and Protein O-Linked Mannose N-Acetylglucosaminyltransferase 1 (β1,2-) (POMGNT1)[192] whereas other groups achieved engineering of defined and elongated O-glycans in CHO[202,203].
Within prokaryotes, the rising bacterial glycoengineering could help to produce vaccine glycoconjugates with immunogenic bacterial glycans as discussed in a previous review[204].

5.3 Carboxylation

Efficient γ-carboxylation of recombinant proteins is another challenge when aiming for human-like PTMs. Similar to N-glycosylation, γ-carboxylation currently requires the usage of mammalian cell platforms. Correct carboxylation is of elementary importance to obtain high yields and functional molecules, especially in the context of vitamin K-dependent coagulation factor production[205]. Most difficult-to-express clotting factors - except for FVIII and Von Willebrand Factor - are carboxylated by the vitamin K-dependent γ-carboxylation mechanism as reviewed by Kumar in 2015[206]. In 2005 Wajih et al. achieved an almost 3-fold increase in carboxylated rhFIX by OE of Vitamin K Epoxide Reductase Complex Subunit 1 (VKORC1), while overexpression of γ-carboxylase lead to the inhibition of functional rhFIX production[207] (Table 1). The study was performed in BHK cells. One year later, the same research group engineered BHK21 cells by siRNA silencing of calumenin, an inhibitor of the γ-carboxylation system, resulting in a downregulation of the gene by ~80%. As an effect, ~80% more functionally active rhFIX was produced by the modified host compared to the untreated control[208]. Although it is known that *Drosophila melanogaster* and S2 cells derived from this insect do express γ-Glutamyl Carboxylase (GGCX) and have a vitamin K processing ability, successful blood coagulation factor manufacturing was not able in these systems yet. Recently, co-expression of mammalian GGCX, VKORC1 and/or Protein Disulfide Isomerase Family A Member 2 (PDIA2) enabled recovery of active rhFVII[209].

5.4 Phosphorylation

Although protein phosphorylation is a very common PTM, it is predominantly associated with intracellular and not with therapeutic proteins[149]. However, proteomic and computational analysis revealed a large number of extracellular phosphorylated proteins including FIX[210,211]. Interestingly, recombinant FIX produced in CHO cells is lacking phosphorylation[212], which does not seem to affect clotting activity. The absence of proof for the importance of phosphorylation of therapeutic proteins

might be one reason why there are so less reports on improving the phosphorylation capacity in industrial expression platforms.

In contrast to mammalian cells, prokaryotes have only a low number of phosphorylated proteins[213,214]. The prokaryotic phosphorylation machinery can be enhanced by overexpressing product-specific protein kinases as presented in *E. coli* by Yue et al. where overexpression of Serine/Threonine-Protein Kinase 1 (SRPK1) increased pre-mRNA-splicing factor SF2 (ASF/SF2) phosphorylation[215] (Table 1). Based on this approach it might also be possible to produce relevant therapeutic proteins with increased phosphorylation in *E. coli* or other expression platforms.

5.5 Sulfation

Recombinant biopharmaceuticals containing sulfation, which are only very few, are generally difficult to express due to molecule and modification complexities[206,216]. Eukaryotic expression platforms are capable to perform this kind of PTM[217]. Heparin is the most famous example for a sulfated difficult-to-express biopharmaceutical and clotting factors rhFVIII and rhFIX have also tyrosine sulfation sites present[206,218]. Currently, Heparin-based drugs are obtained from animal-sources with a contamination crisis in 2008 in the US showing the need for a proper transgenic expression system. Whereas, rhFVIII and rhFIX are mainly produced in CHO, BHK and HEK292 cells with almost fully sulfation[219]. In the study of Datta et al. in 2013, CHO-S cells were engineered to express Golgi-targeted Heparan Sulfate Glucosamine 3-O-Sulfotransferase (HS3ST) 1 resulting in more 2-O, 6-O- and N-sulfo group containing disaccharides in Heparan sulfate (HS), a molecule that is sharing the biosynthesis pathway with Heparin. Thus, the main idea of the researchers was to generate Heparin from HS. Additionally, by the OE of the respective Golgi-bound sulfotransferase an increase of anti-Thrombin activity was achieved at the binding site[220]. Zhang and colleagues overexpressed 2-O-Sulfotransferase (2-OST) as well as C5-epimerase (C_5-epi) in transgenic *E.coli* and used the generated enzymes to produce heparin by bio- and chemically processing of HS[221]. The sulfated glycosaminoglycan chondroitin sulfates, which is used fo the treatment of arthritis, was expressed in metabolically engineered *E.coli* BL21. To achieve the desired tyrosine sulfation, UDP-Glucose-4-Epimerase (kfo) genes A, C and F - enzymes/proteins as part of the capsular polysaccharide production - were transferred from the pathogenic *E.coli* strain K4 into the non-pathogenic strain BL21[222].

5.6 Aggregation & Impurities

5.6.1 Aggregation

Product aggregation can occur at different steps during manufacturing, including fermentation (at intra- and extracellular levels), purification, as well as formulation and storage. The protein aggregates can

lead to adverse side effects after administration[223]. Many efforts therefore have aimed to reduce product aggregation by changing conditions for cultivation[224,225], purification, formulation and storage[226]. A number of genetic engineering approaches decreased intracellular aggregation. In mammalian cells, these strategies include OE of components of the secretion pathway[227], various protein disulfide isomerases and the X-Box-Binding Protein 1 (XBP1)[228-230] or the chaperones Endoplasmic Reticulum Resident Protein 57 (ERp57), Calnexin (CNX) and Calreticulin (CRT)[231,232] which also improve secretion and therefore productivity.

Increasing intracellular chaperone concentration is also a widely used strategy to improve solubility of recombinant proteins in bacterial expression systems. This can be achieved by (i) stimulating bacteria by “heat-shock”, (ii) supplementation of chemical chaperones[233,234] or (iii) OE of certain chaperones (Thiol:Disulfide Interchange Protein (Dsb) A and C) to promote disulfide bond formation and decrease product aggregation[235].

5.6.2 Impurities

During the production of biopharmaceuticals, the cultivation medium might not only contain the recombinant product but also impurities as endogenous host cell proteins (HCP). These HCP must be removed from the therapeutic protein to ensure product quality and prevent adverse health consequences for the patient[81,236–239]. The removal is successful for the majority of impurities during typical downstream purification processes[240].

For mammalian cells, the culture supernatant contains a complex population of HCPs wherein 116 difficult-to-remove HCP were identified by different approaches in CHO cells[241-243]. Based on these target HCPs, Chiu et al. reported the successful CRISPR- and TALEN-guided removal of lipoprotein lipase, a difficult-to-remove HPC[244] (Table 1). Similarly, other HCPs could be KO to generate “cleaner” host cells with decreased protein impurities for improved product quality and simplified downstream purification processes.

Within bacterial expression systems, the contamination with HCPs is not taking place by secretion but during cell lysis to access the recombinant protein. For instance, endotoxins of gram-negative bacteria as *E. coli* need to be removed to ensure the safety of the final product. Similar to the above mentioned example from CHO cells, it is possible to generate bacterial strains lacking certain impurities as presented by a mutated *E. coli* strain with severely low endotoxin levels[245].

5.7 Proteolytic degradation

Proteolytic degradation of biopharmaceuticals induced by host cell-derived proteases is a common problem during bioprocessing occurring throughout all available types of expression systems[246]. Site-specific or -unspecific cleavage of the recombinant product appears intra- and extracellular, respectively, and varies between cell and product types, as well as process conditions[81,246]. Product heterogeneity, clipping of fusion proteins, decreased product titer and quality, a more challenging and complex downstream processing, and potential immunogenicity in patients are the major concerns arising from proteolysis[85]. Site-specific mutagenesis of the recombinant product, changing the host cell system, supplementation with protease inhibitors, and optimization of bioprocessing are options to prevent degradation[81,85]. However, resulting in time-consuming, tedious, and expensive modulation and bioprocessing efforts to in the end obtain a product with characteristics close to the original product[82]. So far, little is known which cell-specific proteases are responsible for product-dependent degradation, thus engineering target selection is more complicated. Nevertheless, a couple of studies are available showing successful engineering of host-specific proteases.

5.7.1 CHO, the main representative of mammalian expression hosts

In 2016, Hu et al. identified Carboxypeptidase D (CpD) as the main source for C-terminal lysine cleavage in a mAb producing CHO cell line[247] (Table 1). By transcription analysis, five carboxypeptidases were investigated in terms of their expression levels in producer cells and their respective parental cell lines. shRNA induced KD of CpD in DUXB-11 and DHFR-positive CHO producers resulted in the absence of C-terminal lysine degradation of the heavy chain. Moreover, high percentage of C-terminal lysine of the mAb were maintained by bi-allelic CpD gene deletion via CRISPR/Cas9. More recently, Laux and colleges achieved decreased non-antibody (Ab) glycoprotein degradation in CHO-K1 cells by TALEN- and ZFN-mediated KO of Matriptase-1 (MT-SP1)[85].

5.7.2 Bacterial expression systems

Endogenous proteases levels are higher in bacterial cytoplasm than in the periplasm of gram-negative bacteria rising the risk for proteolytic degradation of biopharmaceuticals[248]. However, the well established *E. coli* BL21 strain with deletions in two major proteases, the periplasmic outer membrane protease T and the cytoplasmic lon protease, demonstrates the possibility to design bacterial strains with decreased protein degradation[249].

5.7.3 Yeast expression systems

Already in 1998, Kerry-Williams et al. reported reduced proteolytic degradation of recombinant human albumin (rHA) secreted from *S. cerevisiae*[250]. This was reached by classical gene disruption of Yeast Aspartyl Protease 3 (YAP3) alone or in combination with turn-off of Kexin 2 Protease (Kex2p). Wu et al., disrupted encoding Proteinase A (PEP4) and Yapsin Family Member 1 (YPS1) by homologous recombination successfully improving intact hHA and human parathyroid hormone expression levels in *P. pastoris*[251].

5.7.4 Insect-based expression systems

Proteolytic degradation of recombinant therapeutics produced in BICS is often induced by virus-encoded proteases like the Cathepsin L-like Protease (v-cath). By post-transcriptional v-cath silencing in baculovirus-infected Sf-9 cells via double-stranded RNA (dsRNA), Kim et al. achieved an almost 3-fold increase in product yield[63]. Lower level of v-cath-induced proteolysis and a delay in cell lysis were concluded as the reasons for improved productivity. Although green fluorescence protein (GFP) was produced in this study, the engineering strategy can be adapted for a biopharmaceutical producing insect cell line.

5.7.5 Plant-based expression systems

Protein stability and functionality limitations due to degradation, as well as proteolysis during up- and downstream processing in plant-based expression systems are occasions for generally low yields[252]. The high number of proteolytic genes encoded in plant genomes and the lack of proper characterization of those are reasons therefore[53]. Co-expression of tomato cystatin SICYS8, a C1A cysteine-protease inhibitor, along with the murine IgG C5-1 in NT-1 almost doubled the antibody yield in young leaves[253]. By applying asRNA, four endogenous proteases were silenced simultaneously in tobacco BY-2 cells. Hence, Mandal et al., achieved a 4-fold increase in accumulation of a full length IgG1 Ab against a human immunodeficiency virus (HIV)-1 surface protein[254]. Human interleukin (IL)-10 production in transgenic whole-plant *Nicotiana tabacum* (*N. tabacum*) was found to be below the critical expression level as reported by Duwadi and others[255]. Here, transient as well as stable downregulation of the Cysteine Protease 6 (CysP6) by RNAi silencing resulted in accumulation of the total soluble protein fraction in transgenic IL-10 producing tobacco plants.

6. Conclusion

We described the diversity of expression platforms which all represent certain advantages for the production of biopharmaceuticals. However, there is room for optimization since many PTMs are not present in some hosts and because impurities, product aggregation and degradation can have a negative impact on product quality. The availability of sequence information and genetic tools for targeted engineering rapidly speeds up the generation of expression platforms capable to produce biopharmaceuticals with improved product qualities. KD and KO of one or several targets is achieved easier, cheaper and more controlled than ever whereas the OE of transgenes is still a frequent strategy to add certain functions to host cells. Additionally, the possibility to activate silent, or silence active endogenous genes by e.g. CRISPRa/CRISPRi is expanding the genetic toolbox.

Unlike engineering towards improved glycosylation profiles, engineering other quality attributes as carboxylation and phosphorylation are comparably underrepresented with only few targets described. However, novel CRISPR-based screening methods can promote the identification of new targets for improved phenotypes.

The identification and characterization of genes involved in the modification and quality of therapeutic proteins during their production is only one major challenge. The exploration of the relationship between different quality attributes, PTMs as well as the actual *in vivo* function and mode of action of therapeutic proteins is of similar importance. While CHO is superior to the other platforms with regards to most of the PTMs, yeast is catching up with e.g. humanized N-glycosylation and yeast-based therapeutic protein expression can be expected to become a serious contender in any manufacturing strategy[256].

In the authors opinion, there is a trend towards the design of “cleaner” host cells where e.g. the major part of secreted mammalian HCPs are removed to encounter product degradation and complex purification protocols or the “clean-up” of enzyme isoforms maintaining only the one isoform, which is the most important one and capable to retain the function.

Additionally, cell-free systems might enable a fully controlled design of product quality without facing interfering host cell pathways. The cell-free systems can be an alternative to generate therapeutic proteins with homogeneous and improved product quality based on cell lysates from *E. coli* and CHO[257-259].

We conclude that the available genetic engineering tools contribute to a faster and more precise design of expression systems with improved product qualities than ever before. Yet, there is the need for the discovery of novel targets, especially to encounter product degradation and impurities, but also to engineer protein sulfation and phosphorylation.

Table 1. Genetic engineering approaches to improve product quality attributes in different expression platforms (superscript numbers refer to abbreviation index below the table).

Product	Organism	Engineering tool (target)	Described effect	Reference
N-glycosylation				
HCP	CHO (UH)	asRNA KD (CMAH)	Neu5GC ↓	156
EPO	CHO (DuxB11)	CRISPRa (ST6GAL1)	α-2,6-sialylation ↑	157
mAb	CHO, BHK	OE (ST6GAL1)	α-2,6-sialylation ↑	158
mAb	CHO (DG44)	siRNA KD (FUT8)	Fucosylation ↓	159
mAb	CHO (K1)	ZFN KO (FUT8)	Fucosylation ↓	160
HCP	CHO (K1)	CRISPR KO (FUT8)	Fucosylation ↓	161
IFN-ɣ	CHO (DG44)	siRNA/shRNA KD (NEU1, NEU3)	Sialylation ↑	163
EPO, mAb	CHO (S)	CRISPR KO (B4GALT1,B4GALT2,B4GALT3,B4GALT4)	Galactosylation ↓	164
EPO, mAb	CHO (K1)	ZFN KO (19 genes); OE (ST6GAL1)	Fucosylation ↓ α-2,6-sialylation ↑ Galactosylation ↓ Microheterogeneity ↓	165
Vaccine (RSV-F[1])	CHO (DG44)	OE (XYLT)	Xylosylation ↑	167
HCP	*E. coli*	OE (pglB[2]/STT3D)	Complex type ↑	171
scFv	*E. coli*	OE (ALG1,ALG2,ALG13,ALG14)	Complex type ↑	172
HCP	*P. pastoris*	KO (OCH1); OE (GNTI,GNTII,B4GALT,MANI,MANII)	High-mannose ↓	173
mAb	*S. cerevisiae*	KO (ALG3,ALG11,MNN); OE (GNTI,GNTII,STT3D,FLC2[6])	High-mannose ↓	174
Glucose oxidase	*H. polymorpha*[4]	KO (ALG3,ALG11,OCH1); OE (GNTI,GNTII,B4GALT)	High-mannose ↓	175
Lipase 2	*Y. lipolytica*[3]	KO (ALG3); OE (ALG6)	High-mannose ↓	176
GC	*Y. lipolytica*	KO (OCH1,MNN[5]9); OE (MAN)	High-mannose ↓	177
rhEPO	*P. pastoris*	OE of 14 genes, KO of 4 genes	>90% terminal sialylation; Complex type ↑	178
mAb, rhGM-CSF	*P. pastoris*	OE (STT3D)	Macroheterogeneity ↓	179
GC	*N. benthamiana*	RNAi (GNTI)	Immunogenic plant sugar residues ↓	183
mAb	*L. minor*[7]	siRNA KD (XYLT/FUT)	Immunogenic plant sugar residues ↓	184

mAb	*N. tabacum*	CRISPR/Cas9 KO (XYLT/FUT)	Immunogenic plant sugar residues ↓	185
mAb	*N. benthamiana*	RNAi (XYLT/FUT)	Immunogenic plant sugar residues ↓	186
mAb	*B. mori*[8]	OE (MGAT2,B4GALT)	Complex type ↑	187
GST-SfManI, β-Trace[9]	Sf9	OE (MGAT2,B4GALT,ST6GAL1)	Complex type ↑	188
EPO	Sf9	CRISPR/Cas9 KO (FDL)	paucimannose ↓	189
O-glycosylation				
EPO	CHO (K1)	ZFN KO (COSMC, POMGNT1)	O-glycan heterogeneity↓	192
YFP[10], IFNα2B	*N. benthamiana, A. thaliana*[11]	OE (GALNT2[12], WBPP[13])	O-glycosylation ↑	196
YFP, IFNα2B	*N. benthamiana*	OE (GALNT2, GALNT4, WBPP)	O-glycosylation ↑	197
EPO-Fc	*N. benthamiana*	OE (GALNT2, ST3GAL,ST6GAL, B3GALT[14])	di-sialylated core 1 O-glycans ↑	198
IgA1	*N. benthamiana*	KO (XYLT,FUT), OE (GALNT2, C1GALT1)	Core 1 O-glycan ↑	199
podplanin	*S. cerevisiae*	OE (GALE[15], SLC35A2[16], GALNT1, C1GALT1)	Core 1 O-glycan ↑	200
TNFR2[17]-Fc	*P. pastoris*	OE (MAN1B1, POMGNT1, ST6GAL1, B4GALT)	Sialylation ↑	201
PSGL1[18]/mIgG2b	CHO (K1)	OE (B3GNT[19]3, C2GNT[20]1, C3GNT[21]6)	Elongated O-glycans ↑	203
PSGL1/mIgG2b	CHO (K1)	OE (C2GNT1, C3GNT5, C3GNT6, FUT1, FUT2, GALT, GALNT)	matured O-glycans ↑	202
γ-Carboxylation				
rhFIX	BHK	OE (VKORC1)	carboxylation ↑	207
rhFIX	BHK21	siRNA (calumenin)	carboxylation ↑	208
rhFVII	S2	OE (GGCX,VKORC1 and/or PDIA2)	carboxylation ↑	209
Phosphorylation				
ASF/SF2	*E. coli*	OE (SRPK1)	Phosphorylation ↑	215

Sulfation				
HS	CHO-S	OE (Golgi-targeted HS3ST1)	Tyrosine sulfation ↑	220
HS pathway enzymes	*E.coli*	OE (2-OST, C_5-epi)	Recombinant expression of enzymes for downstream application	221
Chondroitin	*E.coli*	OE (kfoA,kfoC,kfoF)	Tyrosine sulfation ↑	222
Aggregation				
mAb	CHO (K1)	OE (SRP14[22])	Intracell. aggregation ↓ Secretion ↑	227
SEAP[23]	CHO (K1)	OE (XBP1)	Secretion ↑	228
mAb	CHO	OE (PDI[24])	Secretion ↑	229
mAb	CHO	OE (PDI, ERO1L[25])	Secretion ↑	230
TPO[26]	CHO (DuxB11)	OE (CNX, CRT)	Productivity ↑	231
TPO	CHO (DuxB11)	OE (ERp57)	Productivity ↑	232
MBP[27]	*E. coli*	OE (Dna[28]K-DnaJ-Grp[29]E,CLPB[30], GroE[31]L-GroES,Ibp[32]A/B)	Intracell. aggregation ↓	233
EGF[33]	*E. coli*	OE (DsbA,DsbC,FKPA[34],SurA[35])	Intracell. aggregation ↓	235
Impurities				
mAb	CHO	KO CRISPR/TALEN	HCP ↓	244
APOA1[36], HSP70[37]	*E. coli*	KO (KDSD[38],GUTQ[39],LPXL[40],LPXM[41],PAGP[42],LPXP[43], EPTA[44])	Endotoxin ↓	245
Product degradation				
GFP	Sf-9	KD dsRNA (v-cath)	Degradation ↓	63
non-Ab glycoprotein	CHO (K1)	KO TALEN/ZFN (MT-SP1)	Degradation ↓	85
mAb	CHO (DUXB-11 and DHFR-positive)	KD/KO shRNA/CRISPR (CpD)	Degradation ↓	247
HCP	*E.coli*	Gene disruption (OmpT, lon)	Degradation ↓	249

rHA	*S. cerevisiae*	Gene disruption (YAP3, Kex2p)	Degradation ↓	250
hHA and human parathyroid hormone	*P. pastoris*	Gene disruption (PEP4, YPS1)	Degradation ↓	251
Ab IgG C5-1	NT-1	OE (SICYS8)	Degradation ↓	253
α-HIV-1 IgG1 Ab	BY-2	KD RNAi (four proteases)	Degradation ↓	254
IL-10	*N. tabacum*	KD RNAi (CysP6)	Degradation ↓	255

[1] RSV-F - respiratory syncytial virus protein F; [2] pglB - Oligosaccharyltransferase pglB; [3] *Y. lipolytica* - *Yarrowia lipolytica*; [4] *H. Polymorpha* - *Hansenula polymorpha*; [5] MNN - α-1,3-Mannosyltransferase; [6] FLC2 - Flavin Carrier Protein 2; [7] *L. minor* - *Lemna minor*; [8] *B.mori* - *Bombyx mori*; [9] GST-SfManI, β-Trace - GST-tagged soluble domain of the Sf class I Golgi mannosidase; [10] YPF - yellow fluorescent protein; [11] *A. thaliana* - *Arabidopsis thaliana*; [12] GALNT2 - Polypeptide N-Acetylgalactosaminyltransferase 2; [13] WBPP - UDP-GlcNAc C4 Epimerase; [14] B3GALT - β-1,3-Galactosyltransferase; [15] GALE - UDP-Galactose-4-Epimerase; [16] SLC35A2 - Solute Carrier Family 35 Member A2; [17] TNFR2 - Tumor Necrosis Factor Receptor 2; [18] PSGL1 - Selectin P ligand; [19] B3GNT - β -1,3-N-acetylglucosaminyltransferase; [20] C2GNT **-** Glucosaminyl (N-Acetyl) Transferase, Core 2; [21] C3GNT **-** Glucosaminyl (N-Acetyl) Transferase, Core 3; [22] SRP14 - Signal Recognition Particle 14; [23] SEAP - secreted embryonic alkaline phosphatase; [24] PDI - Protein Disulfide Isomerase; [25] ERO1L - Endoplasmic Reticulum Oxidoreductase 1 α; [26] TPO - thrombopoietin; [27] MBP - mannose binding protein; [28] Dna - Chaperone protein Dna; [29] Grp - protein Grp; [30] CLPB - Caseinolytic Peptidase B Protein Homolog; [31] GroE - heat shock protein; [32] Ibp - Small Heat Shock Protein; [33] EGF - epidermal growth factor; [34] FKPA - FKBP-Type Peptidyl-Prolyl Cis-Trans Isomerase; [35] SurA - Peptidyl-Prolyl Cis-Trans Isomerase; [36] APOA1 - apolipoprotein A1; [37] HSP70 - heat shock protein 70; [38] KDSD - Arabinose 5-phosphate isomerase KDSD; [39] GUTQ - Arabinose 5-Phosphate Isomerase GUTQ; [40] LPXL - Lipid A Biosynthesis Lauroyltransferase; [41] LPXM - Lipid A Biosynthesis Myristoyltransferase; [42] PAGP - Proliferation-Associated Gene B; [43] LPXP - Lipid A Biosynthesis Palmitoleoyltransferase; [44] EPTA - Phosphoethanolamine Transferase EPTA

Abbreviations:

α, alpha; **β**, beta; **γ**, gamma; **2-OST**, 2-O-Sulfotransferase; ***A. thaliana***, *Arabidopsis thaliana*; **Ab**, antibody; **ADCC**, antibody-dependent cell-mediated cytotoxicity; **ALG,** N-Acetylglucosaminyltransferase; **APOA1**, apolipoprotein A1; **asRNA**, antisense RNA; ***B.mori***, *Bombyx mori*; **B3GALT**, β-1,3-Galactosyltransferase; **B3GNT**, β -1,3-N-acetylglucosaminyltransferase; **B4GALT**, β-1,4-Galactosyltransferase **BHK**, Baby Hamster Kidney; **BICS,** baculovirus-insect cell system; **BY-2,** Bright Yellow 2; **C1GALT1**, Core 1 Synthase Glycoprotein-N-Acetylgalactosamine 3-β-Galactosyltransferase 1; **C2GNT,** Glucosaminyl (N-Acetyl) Transferase, Core 2; **C3GNT,** Glucosaminyl (N-Acetyl) Transferase, Core 3 **C_5-epi**, C5-epimerase; **CAP,** CEVEC's Amniocyte Production; **Cas,** CRISPR-associated; **CHO,** Chinese Hamster Ovary; **CLPB**, Caseinolytic Peptidase B Protein Homolog; **CNX**, Calnexin; **CpD**, Carboxypeptidase D; **CRT**, Calreticulin; **CMAH**, Cytidine Monophospho-N-Acetylneuraminic Acid Hydroxylase; **COSMC**, C1GALT1 Specific Chaperone 1; **CST**, Cerebroside Sulfotransferase; **Cpf1**, CRISPR-associated endonuclease in *Prevotella* and *Francisella*; **CRISPR**, clustered regularly interspaced short palindromic repeats; **crRNA**, CRISPR RNA; **dCas9**, dead Cas9; **CysP6,** Cysteine Protease 6; **Dna**, Chaperone protein Dna; **DSB**, double-strand break; **Dsb**, Thiol:Disulfide Interchange Protein; **dsRNA**, double-stranded RNA; ***E. coli***, *Escherichia coli*; **EGF**, epidermal growth factor; EMA/**EMEA**, European Medicines Agency; **EPTA**, Phosphoethanolamine Transferase EPTA; **ERO1L**, Endoplasmic Reticulum Oxidoreductase 1 α; **ERp57**, Endoplasmic Reticulum Resident Protein 57; **Fc**, fragment crystallizable; **FCS**, Fetal Calf Serum; **FDA**, US Food and Drug Administration; **FKPA**, FKBP-Type Peptidyl-Prolyl Cis-Trans Isomerase **FLC2**, Flavin Carrier Protein 2; **FDL**, *fused lobes* gene; **FUT**, Fucosyltransferase; **GalNAc**, N-Acetylgalactosamine; **GALE**, UDP-Galactose-4-Epimerase; **GALNT2**, Polypeptide N-Acetylgalactosaminyltransferase 2; **GC**, glucocerebrosidase; **GFP**, green fluorescent protein; **GGCX**, γ-Glutamyl Carboxylase; **GM-CSF**, Granulocyte-Macrophage Colony-Stimulating

Factor; **GNE**, Glucosamine (UDP-N-Acetyl)-2-Epimerase; **GNT**, N-acetylglucosaminyltransferase **GOI**, gene of interest; **GroE**, heat shock protein **Grp**, protein Grp; **GST-SfManI, β-Trace**, GST-tagged soluble domain of the Sf class I Golgi mannosidase; **GUTQ**, Arabinose 5-Phosphate Isomerase GUTQ; **gRNA**, guide RNA; ***H. Polymorpha***, *Hansenula polymorpha*; **HCP**, host cell proteins; **HDR**, homology directed repair; **HEK293**, Human Embryonic Kidney 293; **HeLa**, Henrietta Lack's; **HIV**, human immunodeficiency virus; **HS**, Heparan sulfate; **HS3ST**, Heparan Sulfate Glucosamine 3-O-Sulfotransferase; **HSP70**, heat shock protein 70; **Ibp**, Small Heat Shock Protein; **IFN-α**, interferon-α; **IFN-γ**, interferon-γ; **IgA**, Immunoglobulin A; **IgG**, Immunoglobulin G; **IL**, interleukin; **InDel**, insertion/deletion; **KD**, knockdown; **KDSD**, Arabinose 5-Phosphate Isomerase KDSD; **kfo**, UDP-Glucose-4-Epimerase; **KI**, knockin; **KO**, knockout; **Kex2p**, Kexin 2 Protease; ***L. minor***, *Lemna minor*; **LPXL**, Lipid A Biosynthesis Lauroyltransferase; **LPXM**, Lipid A Biosynthesis Myristoyltransferase; **LPXP**, Lipid A Biosynthesis Palmitoleoyltransferase; **mAb**, monoclonal antibody; **MAN**, Mannosidase; **MBP**, mannose binding protein; **MGAT2**, α-1,6-Mannosyl-Glycoprotein β-1,2-N-Acetylglucosaminyltransferase 2; **MNN**, α-1,3-Mannosyltransferase; **MT-SP1**, Membrane Type Serine Protease 1; ***N. benthamiana***, *Nicotiana benthamiana;* ***N. tabacum****, Nicotiana tabacum;* **ncRNA**, non-coding RNA; **NEU**, N-Acetyl-α-Neuraminidase; **Neu5GC**, N-Glycolylneuraminic Acid; **NGS**, next-generation sequencing; **NHEJ**, non-homologous end joining; **NT-1**, *Nicotiana tabacum* 1 cell; **OCH1**, α-1,6-Mannosyltransferase; **OE**, overexpression; **ompT**, Outer Membrane Protease T; ***P. pastoris***, *Pichia pastoris*; **PAGP**, Proliferation-Associated Gene B; **PAM**, protospacer adjacent motif; **PDI**, Protein Disulfide Isomerase; **PDIA2**, Protein Disulfide Isomerase Family A Member 2; **PEP4**, Proteinase A; **pglB**, Oligosaccharyltransferase pglB; **POMGNT1**, Protein O-Linked Mannose N-Acetylglucosaminyltransferase 1 (β-1,2-); **PSGL1**, Selectin P ligand; **PTM**, post-translational modification; **rh**, recombinant human; **rHA**, recombinant human albumin; **rhC1INH**, recombinant human C1 inhibitor; **rhEPO**, recombinant human erythropoietin; **rhF**, recombinant human coagulation factor; **rhG-CSF**, recombinant granulocyte colony-stimulating factor; **rhGH**, recombinant human growth hormone; **RNAi**, RNA interference; **RSV-F**, respiratory syncytial virus protein F; ***S. cerevisiae***, *Saccharomyces cerevisiae*; **scFv**, single-chain variable fragment; **SEAP**, secreted embryonic alkaline phosphatase; **Sf**, *Spodoptera frugiperda*; **sgRNA**, single guide RNA; **shRNA**, small hairpin RNA; **siRNA**, small interfering RNA; **SLC35A2**, Solute Carrier Family 35 Member A2; **SRP14**, Signal Recognition Particle 14; **SRPK1**, Serine/Threonine-Protein Kinase 1; **ST6GAL,** ST6 β-Galactoside α-2,6-Sialyltransferase; **STT3D**, Dolichyl-Diphosphooligosaccharide Protein Glycosyltransferase Subunit STT3; **SurA**, Peptidyl-Prolyl Cis-Trans Isomerase; **TALEN**, transcription activator-like effector nuclease; **TNFR2**, Tumor Necrosis Factor Receptor 2; **TPO**, thrombopoietin; **tracrRNA**, trans-acting RNA; **v-cath**, Cathepsin L-like Protease; **VKORC1**, Vitamin K Epoxide Reductase Complex Subunit 1; **VLPs,** virus-like particles; **WBPP**,UDP-GlcNAc C4 Epimerase; **XBP1**, X-Box-Binding Protein 1; **XYLT,** Xylosyltransferase; ***Y. lipolytica***, *Yarrowia lipolytica;* **YAP3**, Yeast Aspartyl Protease 3; **YFP**, yellow fluorescent protein; **YPS1**, Yapsin Family Member 1; **ZFN**, zinc finger nuclease;

Acknowledgement

T.A and V.S. contributed equally to this review. The Novo Nordisk Foundation (NNF10CC1016517) supported this work. This work has been supported by the Austrian BMWFW, BMVIT, SFG, Standortagentur Tirol, Government of Lower Austria and Business Agency Vienna through the Austrian FFG-COMET- Funding Program. T.A., V.S., N.B., H.F.K. and M.R.A. are receiving funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No. 642663.

Conflict of interest

The authors declare no financial or commercial conflict of interest.

7. References

1. Walsh, G. Biopharmaceutical benchmarks 2014. *Nat. Biotechnol.* **32,** 992–1000 (2014).
2. Demain, A. L. & Vaishnav, P. Production of recombinant proteins by microbes and higher organisms. *Biotechnol. Adv.* **27,** 297–306 (2009).
3. Adrio, J.-L. & Demain, A. L. Recombinant organisms for production of industrial products. *Bioeng. Bugs* **1,** 116–131 (2010).
4. Stavenhagen, K. *et al.* N- and -glycosylation Analysis of Human C1-inhibitor Reveals Extensive Mucin-type -Glycosylation. *Mol. Cell. Proteomics* **17,** 1225–1238 (2018).
5. Lood, C. *et al.* IgG glycan hydrolysis by endoglycosidase S diminishes the proinflammatory properties of immune complexes from patients with systemic lupus erythematosus: a possible new treatment? *Arthritis Rheum.* **64,** 2698–2706 (2012).
6. FDA. Guidance for industry biosimilars: Questions and answers regarding implementation of the biologics price competition and innovation act of 2009. (2012).
7. FDA. Guidance for industry scientific considerations in demonstrating biosimilarity to a reference product. (2012).
8. FDA. Guidance for industry quality considerations in demonstrating biosimilarity to a reference protein product. (2012).
9. FDA. Guidance for industry clinical pharmacology data to support a demonstration of biosimilarity to a reference product. (2014).
10. EMEA. EMEA guideline on similar biological medicinal products containing biotechnology□derived proteins as active substance: Quality issues. (2014).
11. EMEA. Guideline on similar biological medicinal products containing biotechnology□derived proteins as active substance: Non□ clinical and clinical issues. (2014).
12. Schmidt, F. R. Recombinant expression systems in the pharmaceutical industry. *Appl. Microbiol. Biotechnol.* **65,** 363–372 (2004).
13. Karg, S. R. & Kallio, P. T. The production of biopharmaceuticals in plant systems. *Biotechnol. Adv.* **27,** 879–894 (2009).
14. Li, F., Vijayasankaran, N., Shen, A. (yijuan), Kiss, R. & Amanullah, A. Cell culture processes for monoclonal antibody production. *MAbs* **2,** 466 (2010).
15. Dumont, J., Euwart, D., Mei, B., Estes, S. & Kshirsagar, R. Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives. *Crit. Rev. Biotechnol.* **36,** 1110–1122 (2016).
16. Lalonde, M.-E. & Durocher, Y. Therapeutic glycoprotein production in mammalian cells. *J. Biotechnol.* **251,** 128–140 (2017).
17. Casademunt, E. *et al.* The first recombinant human coagulation factor VIII of human origin: human cell line and manufacturing characteristics. *Eur. J. Haematol.* **89,** 165 (2012).
18. Weis, B. L. *et al.* Stable miRNA overexpression in human CAP cells: Engineering alternative production systems for advanced manufacturing of biologics using miR-136 and miR-3074. *Biotechnol. Bioeng.* **115,** 2027–2038 (2018).
19. Xu, X. *et al.* The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. *Nat. Biotechnol.* **29,** 735–741 (2011).
20. Ho, L., Greene, C. L., Schmidt, A. W. & Huang, L. H. Cultivation of HEK 293 cell line and production of a member of the superfamily of G-protein coupled receptors for drug discovery applications using a highly efficient novel bioreactor. *Cytotechnology* **45,** 117–123 (2004).
21. Yao, T. & Asayama, Y. Animal-cell culture media: History, characteristics, and current issues. *Reprod. Med. Biol.* **16,** 99–117 (2017).
22. Subedi, G. P., Johnson, R. W., Moniz, H. A., Moremen, K. W. & Barb, A. High Yield Expression of Recombinant Human Proteins with the Transient Transfection of HEK293 Cells in Suspension. *J. Vis. Exp.* (2015). doi:10.3791/53568
23. Renate Kunert, D. R. Advances in recombinant antibody manufacturing. *Appl. Microbiol. Biotechnol.* **100,** 3451 (2016).
24. Jayapal K.P., Wlaschin K.F., Hu W.S., Yap M.G. Recombinant protein therapeutics from CHO cells—20 years and counting. *Chem. Eng. Prog.* **103,** 40–47 (2007).
25. Vcelar, S. *et al.* Karyotype variation of CHO host cell lines over time in culture characterized by chromosome counting and chromosome painting. *Biotechnol. Bioeng.* **115,** 165–173 (2018).
26. Vcelar, S. *et al.* Changes in Chromosome Counts and Patterns in CHO Cell Lines upon Generation of Recombinant Cell Lines and Subcloning. *Biotechnol. J.* **13,** 1700495 (2018).
27. Wurm, F. M. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat. Biotechnol.* **22,** 1393–1398 (2004).
28. Shimamoto, G., Gegg, C., Boone, T. & Quéva, C. Peptibodies: A flexible alternative format to antibodies. *MAbs* **4,** 586–591 (2012).
29. Baumgarten, T., Jimmy Ytterberg, A., Zubarev, R. A. & de Gier, J.-W. Optimizing Recombinant Protein Production in the Escherichia coli Periplasm Alleviates Stress. *Appl. Environ. Microbiol.* **84,** e00270–18 (2018).
30. Ding, N. *et al.* Increased glycosylation efficiency of recombinant proteins in Escherichia coli by auto-induction. *Biochem. Biophys. Res. Commun.* **485,** 138–143 (2017).
31. Brown, C. W. *et al.* Large-scale analysis of post-translational modifications in E. coli under glucose-limiting conditions. *BMC Genomics* **18,** (2017).
32. Martínez-Alonso, M., González-Montalbán, N., García-Fruitós, E. & Villaverde, A. Learning about protein solubility from bacterial inclusion bodies. *Microb. Cell Fact.* **8,** 4 (2009).
33. Miller, S. I. & Salama, N. R. The gram-negative bacterial periplasm: Size matters. *PLoS Biol.* **16,** e2004935 (2018).
34. Ling Lin Fu *et al.* Protein secretion pathways in Bacillus subtilis: implication for optimization of heterologous protein secretion. *Biotechnol. Adv.* **25,** 1–12 (2007).
35. Sahdev, S., Khattar, S. K. & Saini, K. S. Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Mol. Cell. Biochem.* **307,** 249–264 (2008).
36. Kim, H., Yoo, S. J. & Kang, H. A. Yeast synthetic biology for the production of recombinant therapeutic proteins. *FEMS Yeast Res.* **15,** 1–16 (2015).
37. Irani, Z. A., Kerkhoven, E. J., Shojaosadati, S. A. & Nielsen, J. Genome-scale metabolic model ofPichia pastoriswith native and humanized glycosylation of recombinant proteins. *Biotechnol. Bioeng.* **113,** 961–969 (2015).
38. Bae, J.-H. *et al.* An Efficient Genome-Wide Fusion Partner Screening System for Secretion of Recombinant Proteins in Yeast. *Sci. Rep.* **5,** 12229 (2015).
39. Čiplys, E. *et al.* High-level secretion of native recombinant human calreticulin in yeast. *Microb. Cell Fact.* **14,** 165 (2015).
40. Tang, H. *et al.* N-hypermannose glycosylation disruption enhances recombinant protein production by regulating secretory pathway and cell wall integrity in Saccharomyces cerevisiae. *Sci. Rep.* **6,** 25654 (2016).

41. Mattanovich, D. *et al.* Recombinant Protein Production in Yeasts. in *Methods in Molecular Biology* 329–358 (2011).
42. Cao, J. *et al.* Versatile and on-demand biologics co-production in yeast. *Nat. Commun.* **9,** 77 (2018).
43. Nielsen, J. Production of biopharmaceutical proteins by yeast. *Bioengineered* **4,** 207–211 (2013).
44. Chen, Q. & Davis, K. R. The potential of plants as a system for the development and production of human biologics. *F1000Res.* **5,** (2016).
45. Marsian, J. & Lomonossoff, G. P. Molecular pharming — VLPs made in plants. *Curr. Opin. Biotechnol.* **37,** 201–206 (2016).
46. Reuter, L. J., Bailey, M. J., Joensuu, J. J. & Ritala, A. Scale-up of hydrophobin-assisted recombinant protein production in tobacco BY-2 suspension cells. *Plant Biotechnol. J.* **12,** 402–410 (2014).
47. Raven, N. *et al.* Scaled-up manufacturing of recombinant antibodies produced by plant cells in a 200-L orbitally-shaken disposable bioreactor. *Biotechnol. Bioeng.* **112,** 308–321 (2015).
48. Hellwig, S., Drossard, J., Twyman, R. M. & Fischer, R. Plant cell cultures for the production of recombinant proteins. *Nat. Biotechnol.* **22,** 1415–1422 (2004).
49. Leuzinger, K. *et al.* Efficient Agroinfiltration of Plants for High-level Transient Expression of Recombinant Proteins. *J. Vis. Exp.* (2013). doi:10.3791/50521
50. Zimran, A. *et al.* Pivotal trial with plant cell-expressed recombinant glucocerebrosidase, taliglucerase alfa, a novel enzyme replacement therapy for Gaucher disease. *Blood* **118,** 5767–5773 (2011).
51. Mercx, S. *et al.* Inactivation of the β(1,2)-xylosyltransferase and the α(1,3)-fucosyltransferase genes in Nicotiana tabacum BY-2 Cells by a Multiplex CRISPR/Cas9 Strategy Results in Glycoproteins without Plant-Specific Glycans. *Front. Plant Sci.* **8,** (2017).
52. Strasser, R., Altmann, F. & Steinkellner, H. Controlled glycosylation of plant-produced recombinant proteins. *Curr. Opin. Biotechnol.* **30,** 95–100 (2014).
53. Pillay, P., Schlüter, U., van Wyk, S., Kunert, K. J. & Vorster, B. J. Proteolysis of recombinant proteins in bioengineered plant cells. *Bioengineered* **5,** 15–20 (2014).
54. Mandal, M. K., Ahvari, H., Schillberg, S. & Schiermeyer, A. Tackling Unwanted Proteolysis in Plant Production Hosts Used for Molecular Farming. *Front. Plant Sci.* **7,** 267 (2016).
55. Maghodia, A. B., Geisler, C. & Jarvis, D. L. Characterization of an Sf-rhabdovirus-negative Spodoptera frugiperda cell line as an alternative host for recombinant protein production in the baculovirus-insect cell system. *Protein Expr. Purif.* **122,** 45–55 (2016).
56. Cox, M. M. J. Recombinant protein vaccines produced in insect cells. *Vaccine* **30,** 1759–1766 (2012).
57. Roldão, A., Vicente, T., Peixoto, C., Carrondo, M. J. T. & Alves, P. M. Quality control and analytical methods for baculovirus-based products. *J. Invertebr. Pathol.* **107 Suppl,** S94–105 (2011).
58. Chavez-Pena, C. & Kamen, A. A. RNA interference technology to improve the baculovirus-insect cell expression system. *Biotechnol. Adv.* **36,** 443–451 (2018).
59. Meghrous, J. *et al.* Dissolved carbon dioxide determines the productivity of a recombinant hemagglutinin component of an influenza vaccine produced by insect cells. *Biotechnol. Bioeng.* **112,** 2267–2275 (2015).
60. Mbewana, S., Mortimer, E., Pêra, F. F. P. G., Hitzeroth, I. I. & Rybicki, E. P. Production of H5N1 Influenza Virus Matrix Protein 2 Ectodomain Protein Bodies in Tobacco Plants and in Insect Cells as a Candidate Universal Influenza Vaccine. *Front Bioeng Biotechnol* **3,** 197 (2015).
61. Ailor, E. & Betenbaugh, M. J. Modifying secretion and post-translational processing in insect cells. *Curr. Opin. Biotechnol.* **10,** 142–145 (1999).
62. Geisler, C., Aumiller, J. J. & Jarvis, D. L. A fused lobes gene encodes the processing beta-N-acetylglucosaminidase in Sf9 cells. *J. Biol. Chem.* **283,** 11330–11339 (2008).
63. Kim, E. J., Kramer, S. F., Hebert, C. G., Valdes, J. J. & Bentley, W. E. Metabolic engineering of the baculovirus-expression system via inverse 'shotgun' genomic analysis and RNA interference (dsRNA) increases product yield and cell longevity. *Biotechnol. Bioeng.* **98,** 645–654 (2007).
64. Bönig, H. *et al.* Glycosylated vs non-glycosylated granulocyte colony-stimulating factor (G-CSF)--results of a prospective randomised monocentre study. *Bone Marrow Transplant.* **28,** 259–264 (2001).
65. Adolf, G. R., Kalsner, I., Ahorn, H., Maurer-Fogy, I. & Cantell, K. Natural human interferon-α2 isO-glycosylated. *Biochem. J* **276,** 511–518 (1991).
66. van Veen, H. A. *et al.* Characterization of recombinant human C1 inhibitor secreted in milk of transgenic rabbits. *J. Biotechnol.* **162,** 319–326 (2012).
67. Zhang, L., Luo, S. & Zhang, B. The use of lectin microarray for assessing glycosylation of therapeutic proteins. *MAbs* **8,** 524–535 (2016).
68. Houel, S. *et al.* N- and O-glycosylation analysis of etanercept using liquid chromatography and quadrupole time-of-flight mass spectrometry equipped with electron-transfer dissociation functionality. *Anal. Chem.* **86,** 576–584 (2014).
69. Lai, J. D. *et al.* N-linked glycosylation modulates the immunogenicity of recombinant human factor VIII in hemophilia A mice. *Haematologica* (2018). doi:10.3324/haematol.2018.188219
70. European Medicines Agency. Annex to guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: Non-clinical and clinical issues. Guidance on similar medicinal products containing recombinant human insulin. (2006).
71. Liu, L. Antibody glycosylation and its impact on the pharmacokinetics and pharmacodynamics of monoclonal antibodies and Fc-fusion proteins. *J. Pharm. Sci.* **104,** 1866–1884 (2015).
72. Shaaltiel, Y. *et al.* Production of glucocerebrosidase with terminal mannose glycans for enzyme replacement therapy of Gaucher's disease using a plant cell system. *Plant Biotechnol. J.* **5,** 579–590 (2007).
73. Berkner, K. L. Vitamin K-dependent carboxylation. *Vitam. Horm.* **78,** 131–156 (2008).
74. Walsh, G. *Post-translational Modification of Protein Biopharmaceuticals*. (John Wiley & Sons, 2009).
75. Hansson, K. & Stenflo, J. Post-translational modifications in proteins involved in blood coagulation. *J. Thromb. Haemost.* **3,** 2633–2648 (2005).
76. Kaufman, R. J. Post-translational modifications required for coagulation factor secretion and function. *Thromb. Haemost.* **79,** 1068–1079 (1998).
77. Moore, K. L. The biology and enzymology of protein tyrosine O-sulfation. *J. Biol. Chem.* **278,** 24243–24246 (2003).
78. Ezban, M., Vad, K. & Kjalke, M. Turoctocog alfa (NovoEight®)--from design to clinical proof of concept. *Eur. J. Haematol.* **93,** 369–

376 (2014).
79. Lebendiker, M. & Danieli, T. Production of prone-to-aggregate proteins. *FEBS Lett.* **588,** 236–246 (2014).
80. Dorai, H. *et al.* Characterization of the proteases involved in the N-terminal clipping of glucagon-like-peptide-1-antibody fusion proteins. *Biotechnol. Prog.* **27,** 220–231 (2011).
81. Gao, S. X. *et al.* Fragmentation of a highly purified monoclonal antibody attributed to residual CHO cell protease activity. *Biotechnol. Bioeng.* **108,** 977–982 (2010).
82. Robert, F. *et al.* Degradation of an Fc-fusion recombinant protein by host cell proteases: Identification of a CHO cathepsin D protease. *Biotechnol. Bioeng.* **104,** 1132–1141 (2009).
83. Hansen, K., Kjalke, M., Rasmussen, P. B., Kongerslev, L. & Ezban, M. Proteolytic cleavage of recombinant two-chain factor VIII during cell culture production is mediated by protease(s) from lysed cells. The use of pulse labelling directly in production medium. *Cytotechnology* **24,** 227–234 (1997).
84. Rimbon, J., Sánchez-Kopper, A., Wahl, A. & Takors, R. Monitoring intracellular protein degradation in antibody-producing Chinese hamster ovary cells. *Eng. Life Sci.* **15,** 499–508 (2015).
85. Laux, H. *et al.* Degradation of recombinant proteins by Chinese hamster ovary host cell proteases is prevented by matriptase-1 knockout. *Biotechnol. Bioeng.* (2018). doi:10.1002/bit.26731
86. Blattner, F. R. The Complete Genome Sequence of Escherichia coli K-12. *Science* **277,** 1453–1462 (1997).
87. Kim, M., Rai, N., Zorraquino, V. & Tagkopoulos, I. Multi-omics integration accurately predicts cellular state in unexplored conditions for Escherichia coli. *Nat. Commun.* **7,** 13090 (2016).
88. Nandakumar, S., Ma, H. & Khan, A. S. Whole-Genome Sequence of the Spodoptera frugiperda Sf9 Insect Cell Line. *Genome Announc.* **5,** (2017).
89. Sturmberger, L. *et al.* Refined Pichia pastoris reference genome sequence. *J. Biotechnol.* **235,** 121–131 (2016).
90. Rupp, O. *et al.* A reference genome of the Chinese hamster based on a hybrid assembly strategy. *Biotechnol. Bioeng.* **115,** 2087–2100 (2018).
91. Jin, J. *et al.* Integrated transcriptomics and metabolomics analysis to characterize cold stress responses in Nicotiana tabacum. *BMC Genomics* **18,** 496 (2017).
92. Goodwin, S., McPherson, J. D. & McCombie, W. R. Coming of age: ten years of next-generation sequencing technologies. *Nat. Rev. Genet.* **17,** 333–351 (2016).
93. Stolfa, G. *et al.* CHO-Omics Review: The Impact of Current and Emerging Technologies on Chinese Hamster Ovary Based Bioproduction. *Biotechnol. J.* **13,** e1700227 (2018).
94. Brunk, E. *et al.* Characterizing Strain Variation in Engineered E. coli Using a Multi-Omics-Based Workflow. *Cell Syst* **2,** 335–346 (2016).
95. Feichtinger, J. *et al.* Comprehensive genome and epigenome characterization of CHO cells in response to evolutionary pressures and over time. *Biotechnol. Bioeng.* **113,** 2241–2253 (2016).
96. Shcherbakova, O. G., Lanzov, V. A., Ogawa, H. & Filatov, M. V. Overexpression of bacterial RecA protein stimulates homologous recombination in somatic mammalian cells. *Mutat. Res.* **459,** 65–71 (2000).
97. Chen, X. *et al.* Overexpression of bacterial ethylene-forming enzyme gene in Trichoderma reesei enhanced the production of ethylene. *Int. J. Biol. Sci.* **6,** 96–106 (2010).
98. Zhao, W.-N. & McAlister-Henn, L. Expression and Gene Disruption Analysis of the Isocitrate Dehydrogenase Family in Yeast†. *Biochemistry* **35,** 7873–7878 (1996).
99. Lee, M. E., Aswani, A., Han, A. S., Tomlin, C. J. & Dueber, J. E. Expression-level optimization of a multi-enzyme pathway in the absence of a high-throughput assay. *Nucleic Acids Res.* **41,** 10668–10678 (2013).
100. Lewis, A. M., Abu-Absi, N. R., Borys, M. C. & Li, Z. J. The use of 'Omics technology to rationally improve industrial mammalian cell line performance. *Biotechnol. Bioeng.* **113,** 26–38 (2015).
101. Michael Snyder, J. E. G. G. Systems Biology from a Yeast Omics Perspective. *FEBS Lett.* **583,** 3895 (2009).
102. Kuystermans, D., Krampe, B., Swiderek, H. & Al-Rubeai, M. Using cell engineering and omic tools for the improvement of cell culture processes. *Cytotechnology* **53,** 3 (2007).
103. Gutierrez, J. M. & Lewis, N. E. Optimizing eukaryotic cell hosts for protein production through systems biotechnology and genome-scale modeling. *Biotechnol. J.* **10,** 939–949 (2015).
104. Nguyen, Q., Nielsen, L. K. & Reid, S. Genome Scale Transcriptomics of Baculovirus-Insect Interactions. *Viruses* **5,** 2721 (2013).
105. Gupta, S. K. & Shukla, P. Gene editing for cell engineering: trends and applications. *Crit. Rev. Biotechnol.* **37,** 672–684 (2017).
106. Fischer, S. *et al.* A functional high-content miRNA screen identifies miR-30 family to boost recombinant protein production in CHO cells. *Biotechnol. J.* **9,** 1279–1292 (2014).
107. Wu C, E. al. Engineering Periodic shRNA for Enhanced Silencing Efficacy. - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/27053374. (Accessed: 27th August 2018)
108. Kseniya Gavrilov, W. M. S. Therapeutic siRNA: Principles, Challenges, and Strategies. *Yale J. Biol. Med.* **85,** 187 (2012).
109. Inwood, S., Betenbaugh, M. J. & Shiloach, J. Methods for Using Small Non-Coding RNAs to Improve Recombinant Protein Expression in Mammalian Cells. *Genes* **9,** (2018).
110. Agrawal, N. *et al.* RNA Interference: Biology, Mechanism, and Applications. *Microbiol. Mol. Biol. Rev.* **67,** 657 (2003).
111. Lam, J. K. W., Chow, M. Y. T., Zhang, Y. & Leung, S. W. S. siRNA Versus miRNA as Therapeutics for Gene Silencing. *Mol. Ther. Nucleic Acids* **4,** e252 (2015).
112. Kim, H. J., Lee, H. J., Kim, H., Cho, S. W. & Kim, J.-S. Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Res.* **19,** 1279–1288 (2009).
113. Miller, J. C. *et al.* A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* **29,** 143–148 (2011).
114. Epinat, J.-C. *et al.* A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. *Nucleic Acids Res.* **31,** 2952–2962 (2003).
115. Anthony J. Davis, D. J. C. DNA double strand break repair via non-homologous end-joining. *Transl. Cancer Res.* **2,** 130 (2013).
116. Liang, F., Han, M., Romanienko, P. J. & Jasin, M. Homology-directed repair is a major double-strand break repair pathway in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 5172 (1998).
117. Santiago, Y. *et al.* Targeted gene knockout in mammalian cells by using engineered zinc-finger nucleases. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 5809–5814 (2008).

118. Sakuma, T. *et al.* Homologous Recombination-Independent Large Gene Cassette Knock-in in CHO Cells Using TALEN and MMEJ-Directed Donor Plasmids. *Int. J. Mol. Sci.* **16,** 23849–23866 (2015).
119. Heyer, W.-D., Ehmsen, K. T. & Liu, J. Regulation of homologous recombination in eukaryotes. *Annu. Rev. Genet.* **44,** 113–139 (2010).
120. Miyaoka, Y. *et al.* Systematic quantification of HDR and NHEJ reveals effects of locus, nuclease, and cell type on genome-editing. *Sci. Rep.* **6,** 23549 (2016).
121. Mali, P. *et al.* RNA-Guided Human Genome Engineering via Cas9. *Science* **339,** 823 (2013).
122. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337,** 816–821 (2012).
123. Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507,** 62–67 (2014).
124. Karvelis, T. *et al.* crRNA and tracrRNA guide Cas9-mediated DNA interference in Streptococcus thermophilus. *RNA Biol.* **10,** 841 (2013).
125. Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M. & Joung, J. K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* **32,** 279–284 (2014).
126. Ul Ain, Q., Chung, J. Y. & Kim, Y.-H. Current and future delivery systems for engineered nucleases: ZFN, TALEN and RGEN. *J. Control. Release* **205,** 120–127 (2015).
127. Schmieder, V. *et al.* Enhanced Genome Editing Tools For Multi-Gene Deletion Knock-Out Approaches Using Paired CRISPR sgRNAs in CHO Cells. *Biotechnol. J.* **13,** 1700211 (2017).
128. Bydlinski, N. *et al.* The contributions of individual galactosyltransferases to protein specific N-glycan processing in Chinese Hamster Ovary cells. *J. Biotechnol.* **282,** 101–110 (2018).
129. Zheng, Q. *et al.* Precise gene deletion and replacement using the CRISPR/Cas9 system in human cells. *Biotechniques* **57,** 115–124 (2014).
130. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **152,** 1173 (2013).
131. O'Geen, H. *et al.* dCas9-based epigenome editing suggests acquisition of histone methylation is not sufficient for target gene repression. *Nucleic Acids Res.* **45,** 9901 (2017).
132. Shawn Liu, X. *et al.* Editing DNA methylation in the mammalian genome. *Cell* **167,** 233 (2016).
133. Xu, X. *et al.* A CRISPR-based approach for targeted DNA demethylation. *Cell Discov* **2,** 16009 (2016).
134. Verwaal, R., Buiting-Wiessenhaan, N., Dalhuijsen, S. & Roubos, J. A. CRISPR/Cpf1 enables fast and simple genome editing of Saccharomyces cerevisiae. *Yeast* **35,** 201–211 (2018).
135. Zaidi, S. S.-E.-A., Mahfouz, M. M. & Mansoor, S. CRISPR-Cpf1: A New Tool for Plant Genome Editing. *Trends Plant Sci.* **22,** 550–553 (2017).
136. Lee, J. S., Grav, L. M., Pedersen, L. E., Lee, G. M. & Kildegaard, H. F. Accelerated homology-directed targeted integration of transgenes in Chinese hamster ovary cells via CRISPR/Cas9 and fluorescent enrichment. *Biotechnol. Bioeng.* **113,** 2518–2523 (2016).
137. Eisenhut, P. *et al.* A CRISPR/Cas9 based engineering strategy for overexpression of multiple genes in Chinese hamster ovary cells. *Metab. Eng.* **48,** 72–81 (2018).
138. Shalem, O., Sanjana, N. E. & Zhang, F. High-throughput functional genomics using CRISPR–Cas9. *Nat. Rev. Genet.* **16,** 299–311 (2015).
139. Arroyo, J. D. *et al.* A Genome-wide CRISPR Death Screen Identifies Genes Essential for Oxidative Phosphorylation. *Cell Metab.* **24,** 875–885 (2016).
140. Birsoy, K. *et al.* An Essential Role of the Mitochondrial Electron Transport Chain in Cell Proliferation Is to Enable Aspartate Synthesis. *Cell* **162,** 540–551 (2015).
141. Zhu, S. *et al.* Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR-Cas9 library. *Nat. Biotechnol.* **34,** 1279–1286 (2016).
142. Liu, X. S. *et al.* Editing DNA Methylation in the Mammalian Genome. *Cell* **167,** 233–247.e17 (2016).
143. Kweon, J. & Kim, Y. High-throughput genetic screens using CRISPR–Cas9 system. *Arch. Pharm. Res.* (2018). doi:10.1007/s12272-018-1029-z
144. Kim, H. S. *et al.* Arrayed CRISPR screen with image-based assay reliably uncovers host genes required for coxsackievirus infection. *Genome Res.* **28,** 859–868 (2018).
145. Shalem, O. *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343,** 84–87 (2014).
146. Agrotis, A. & Ketteler, R. A new age in functional genomics using CRISPR/Cas9 in arrayed library screening. *Front. Genet.* **6,** 300 (2015).
147. Fan, Y., Kildegaard, H. F. & Andersen, M. R. Engineer Medium and Feed for Modulating N-Glycosylation of Recombinant Protein Production in CHO Cell Culture. in *Methods in Molecular Biology* 209–226 (2017).
148. Friedman, B. *et al.* A comparison of the pharmacological properties of carbohydrate remodeled recombinant and placental-derived beta-glucocerebrosidase: implications for clinical efficacy in treatment of Gaucher disease. *Blood* **93,** 2807–2816 (1999).
149. Walsh, G. Post-translational modifications of protein biopharmaceuticals. *Drug Discov. Today* **15,** 773–780 (2010).
150. Lis, H., Halina, L. I. S. & Sharon, N. Protein glycosylation. Structural and functional aspects. *Eur. J. Biochem.* **218,** 1–27 (1993).
151. Werner, R. G., Kopp, K. & Schlueter, M. Glycosylation of therapeutic proteins in different production systems. *Acta Paediatr.* **96,** 17–22 (2007).
152. Goochee, C. F. & Monica, T. Environmental Effects on Protein Glycosylation. *Nat. Biotechnol.* **8,** 421–427 (1990).
153. Harrison, R. L. & Jarvis, D. L. Protein N-glycosylation in the baculovirus-insect cell expression system and engineering of insect cells to produce 'mammalianized' recombinant glycoproteins. *Adv. Virus Res.* **68,** 159–191 (2006).
154. Tanner, W. & Lehle, L. Protein glycosylation in yeast. *Biochim. Biophys. Acta* **906,** 81–99 (1987).
155. Sareneva, T., Cantell, K., Pyhälä, L., Pirhonen, J. & Julkunen, I. Effect of Carbohydrates on the Pharmacokinetics of Human Interferon-γ. *J. Interferon Res.* **13,** 267–269 (1993).
156. Chenu, S. *et al.* Reduction of CMP-N-acetylneuraminic acid hydroxylase activity in engineered Chinese hamster ovary cells using an antisense-RNA strategy. *Biochim. Biophys. Acta* **1622,** 133–144 (2003).
157. Marx, N. *et al.* CRISPR-Based Targeted Epigenetic Editing Enables Gene Expression Modulation of the Silenced Beta-Galactoside

Alpha-2,6-Sialyltransferase 1 in CHO Cells. *Biotechnol. J.* e1700217 (2018).
158. Raymond, C. *et al.* Production of Highly Sialylated Monoclonal Antibodies. in *Glycosylation* (2012).
159. Mori, K. *et al.* Engineering Chinese hamster ovary cells to maximize effector function of produced antibodies using FUT8 siRNA. *Biotechnol. Bioeng.* **88,** 901–908 (2004).
160. Malphettes, L. *et al.* Highly efficient deletion of FUT8 in CHO cell lines using zinc-finger nucleases yields cells that produce completely nonfucosylated antibodies. *Biotechnol. Bioeng.* **106,** 774–783 (2010).
161. Ronda, C. *et al.* Accelerating genome editing in CHO cells using CRISPR Cas9 and CRISPy, a web-based target finding tool. *Biotechnol. Bioeng.* **111,** 1604–1616 (2014).
162. Monroe, R. S. & Huber, B. E. The major form of the murine asialoglycoprotein receptor: cDNA sequence and expression in liver, testis and epididymis [Gene 148 (1995) 237–244]. *Gene* **161,** 307 (1995).
163. Zhang, M., Koskie, K., Ross, J. S., Kayser, K. J. & Caple, M. V. Enhancing glycoprotein sialylation by targeted gene silencing in mammalian cells. *Biotechnol. Bioeng.* **105,** 1094–1105 (2010).
164. Amann, T. *et al.* CRISPR/Cas9-Multiplexed Editing of Chinese Hamster Ovary B4Gal-T1, 2, 3, and 4 Tailors N -Glycan Profiles of Therapeutics and Secreted Host Cell Proteins. *Biotechnol. J.* (2018). doi:10.1002/biot.201800111
165. Yang, Z. *et al.* Engineered CHO cells for production of diverse, homogeneous glycoproteins. *Nat. Biotechnol.* **33,** 842–844 (2015).
166. Tejwani, V., Andersen, M. R., Nam, J. H. & Sharfstein, S. T. Glycoengineering in CHO Cells: Advances in Systems Biology. *Biotechnol. J.* **13,** 1700234 (2018).
167. Sandig, G. *et al.* Engineering of CHO Cells for the Production of Recombinant Glycoprotein Vaccines with Xylosylated N-glycans. *Bioengineering (Basel)* **4,** (2017).
168. Szymanski, C. M., Yao, R., Ewing, C. P., Trust, T. J. & Guerry, P. Evidence for a system of general protein glycosylation in Campylobacter jejuni. *Mol. Microbiol.* **32,** 1022–1030 (1999).
169. Castric, P. pilO, a gene required for glycosylation of Pseudomonas aeruginosa 1244 pilin. *Microbiology* **141 (Pt 5),** 1247–1254 (1995).
170. Thibault, P. *et al.* Identification of the carbohydrate moieties and glycosylation motifs in Campylobacter jejuni flagellin. *J. Biol. Chem.* **276,** 34862–34870 (2001).
171. Wacker, M. N-Linked Glycosylation in Campylobacter jejuni and Its Functional Transfer into E. coli. *Science* **298,** 1790–1793 (2002).
172. Valderrama-Rincon, J. D. *et al.* An engineered eukaryotic protein glycosylation pathway in Escherichia coli. *Nat. Chem. Biol.* **8,** 434–436 (2012).
173. Jacobs, P. P., Geysens, S., Vervecken, W., Contreras, R. & Callewaert, N. Engineering complex-type N-glycosylation in Pichia pastoris using GlycoSwitch technology. *Nat. Protoc.* **4,** 58–70 (2009).
174. Nasab, F. P., Aebi, M., Bernhard, G. & Frey, A. D. A Combined System for Engineering Glycosylation Efficiency and Glycan Structure in Saccharomyces cerevisiae. *Appl. Environ. Microbiol.* **79,** 997 (2013).
175. Wang, H., Song, H.-L., Wang, Q. & Qiu, B.-S. Expression of glycoproteins bearing complex human-like glycans with galactose terminal in Hansenula polymorpha. *World J. Microbiol. Biotechnol.* **29,** 447–458 (2013).
176. De Pourcq, K. *et al.* Engineering Yarrowia lipolytica to produce glycoproteins homogeneously modified with the universal Man3GlcNAc2 N-glycan core. *PLoS One* **7,** e39976 (2012).
177. De Pourcq, K. *et al.* Engineering the yeast Yarrowia lipolytica for the production of therapeutic proteins homogeneously glycosylated with Man8GlcNAc2 and Man5GlcNAc2. *Microb. Cell Fact.* **11,** 53 (2012).
178. Hamilton, S. R. *et al.* Humanization of yeast to produce complex terminally sialylated glycoproteins. *Science* **313,** 1441–1443 (2006).
179. Choi BK, E. al. Improvement of N-glycan site occupancy of therapeutic glycoproteins produced in Pichia pastoris. - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/22569635. (Accessed: 22nd August 2018)
180. Altmann, F. The role of protein glycosylation in allergy. - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/17033195. (Accessed: 22nd August 2018)
181. Shaaltiel Y, E. al. Production of glucocerebrosidase with terminal mannose glycans for enzyme replacement therapy of Gaucher's disease using a plant cell system. - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/17524049. (Accessed: 22nd August 2018)
182. Tekoah, Y. *et al.* Glycosylation and functionality of recombinant β-glucocerebrosidase from various production systems. *Biosci. Rep.* **33,** (2013).
183. Limkul, J. *et al.* The production of human glucocerebrosidase in glyco-engineered Nicotiana benthamiana plants. *Plant Biotechnol. J.* **14,** 1682–1694 (2016).
184. Cox, K. M. *et al.* Glycan optimization of a human monoclonal antibody in the aquatic plant Lemna minor. *Nat. Biotechnol.* **24,** 1591–1597 (2006).
185. Mercx S, E. al. Inactivation of the β(1,2)-xylosyltransferase and the α(1,3)-fucosyltransferase genes in Nicotiana tabacum BY-2 Cells by a Multiplex CRISPR/Cas9 St... - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/28396675. (Accessed: 22nd August 2018)
186. Strasser R, E. al. Generation of glyco-engineered Nicotiana benthamiana for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure. - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/18346095. (Accessed: 22nd August 2018)
187. Kato, T. *et al.* N-Glycan Modification of a Recombinant Protein via Coexpression of Human Glycosyltransferases in Silkworm Pupae. *Sci. Rep.* **7,** 1409 (2017).
188. Hollister, J., Grabenhorst, E., Nimtz, M., Conradt, H. & Jarvis, D. L. Engineering the protein N-glycosylation pathway in insect cells for production of biantennary, complex N-glycans. *Biochemistry* **41,** 15093–15104 (2002).
189. Mabashi-Asazuma, H. & Jarvis, D. L. CRISPR-Cas9 vectors for genome editing and host engineering in the baculovirus–insect cell system. *Proceedings of the National Academy of Sciences* **114,** 9068–9073 (2017).
190. Bennett, E. P. *et al.* Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. *Glycobiology* **22,** 736–756 (2012).
191. Dicker, M. & Strasser, R. Using glyco-engineering to produce therapeutic proteins. *Expert Opin. Biol. Ther.* **15,** 1501–1516 (2015).
192. Yang, Z. *et al.* The GalNAc-type O-Glycoproteome of CHO Cells Characterized by the SimpleCell Strategy. *Mol. Cell. Proteomics* **13,** 3224–3235 (2014).
193. Taschwer, M. *et al.* Growth, productivity and protein glycosylation in a CHO EpoFc producer cell line adapted to glutamine-free

growth. *J. Biotechnol.* **157,** 295–303 (2012).
194. Kong, Y. *et al.* Probing polypeptide GalNAc-transferase isoform substrate specificities by in vitro analysis. *Glycobiology* **25,** 55–65 (2015).
195. Hermeling, S., Crommelin, D. J. A., Schellekens, H. & Jiskoot, W. Structure-immunogenicity relationships of therapeutic proteins. *Pharm. Res.* **21,** 897–903 (2004).
196. Yang, Z. *et al.* Toward stable genetic engineering of human O-glycosylation in plants. *Plant Physiol.* **160,** 450–463 (2012).
197. Yang Z, E. al. Toward stable genetic engineering of human O-glycosylation in plants. - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/22791304. (Accessed: 22nd August 2018)
198. Castilho, A. *et al.* Engineering of Sialylated Mucin-typeO-Glycosylation in Plants. *J. Biol. Chem.* **287,** 36518–36526 (2012).
199. Dicker M, E. al. Transient Glyco-Engineering to Produce Recombinant IgA1 with Defined N- and O-Glycans in Plants. - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/26858738. (Accessed: 22nd August 2018)
200. Amano, K. *et al.* Engineering of mucin-type human glycoproteins in yeast cells. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 3232–3237 (2008).
201. Hamilton, S. R. *et al.* Production of sialylated O-linked glycans in Pichia pastoris. *Glycobiology* **23,** 1192–1203 (2013).
202. Lindberg, L. *et al.* Mucin-type fusion proteins with blood group A or B determinants on defined O-glycan core chains produced in glycoengineered Chinese hamster ovary cells and their use as immunoaffinity matrices. *Glycobiology* **23,** 720–735 (2013).
203. Liu, J., Jin, C., Cherian, R. M., Karlsson, N. G. & Holgersson, J. O -glycan repertoires on a mucin-type reporter protein expressed in CHO cell pools transiently transfected with O -glycan core enzyme cDNAs. *J. Biotechnol.* **199,** 77–89 (2015).
204. Yates, L. E., Mills, D. C. & DeLisa, M. P. Bacterial Glycoengineering as a Biosynthetic Route to Customized Glycomolecules. in *Advances in Biochemical Engineering/Biotechnology* (2018).
205. Vatandoost, J. & Pakdaman, S. F. The Effects of Influencing Factors on γ-carboxylation and Expression of Recombinant Vitamin K Dependent Coagulation Factors. *J. Biomed. Biotechnol.* **1,** (2016).
206. Kumar, S. R. Industrial production of clotting factors: Challenges of expression, and choice of host cells. *Biotechnol. J.* **10,** 995–1004 (2015).
207. Wajih, N., Hutson, S. M., Owen, J. & Wallin, R. Increased production of functional recombinant human clotting factor IX by baby hamster kidney cells engineered to overexpress VKORC1, the vitamin K 2,3-epoxide-reducing enzyme of the vitamin K cycle. *J. Biol. Chem.* **280,** 31603–31607 (2005).
208. Wajih, N., Hutson, S. M. & Wallin, R. siRNA silencing of calumenin enhances functional factor IX production. *Blood* **108,** 3757 (2006).
209. Nagahashi, K., Umemura, K., Kanayama, N. & Iwaki, T. Successful synthesis of active human coagulation factor VII by co-expression of mammalian gamma-glutamyl carboxylase and modification of vit.K cycle in Drosophila Schneider S2 cells. *Cytotechnology* **69,** 317 (2017).
210. Kaushik P, E. al. The expression pattern of the phosphoproteome is significantly changed during the growth phases of recombinant CHO cell culture. - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/30076757. (Accessed: 22nd August 2018)
211. Yalak, G. & Vogel, V. Extracellular Phosphorylation and Phosphorylated Proteins: Not Just Curiosities But Physiologically Important. *Sci. Signal.* **5,** re7–re7 (2012).
212. Rand, M. D., Kalafatis, M. & Mann, K. G. Platelet coagulation factor Va: the major secretory platelet phosphoprotein. *Blood* **83,** 2180–2190 (1994).
213. Enami, M. & Ishihama, A. Protein phosphorylation in Escherichia coli and purification of a protein kinase. *J. Biol. Chem.* **259,** 526–533 (1984).
214. Macek B, E. al. Phosphoproteome analysis of E. coli reveals evolutionary conservation of bacterial Ser/Thr/Tyr phosphorylation. - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/17938405. (Accessed: 22nd August 2018)
215. Yue, B.-G., Ajuh, P., Akusjärvi, G., Lamond, A. I. & Kreivi, J.-P. Functional coexpression of serine protein kinase SRPK1 and its substrate ASF/SF2 in Escherichia coli. *Nucleic Acids Res.* **28,** e14 (2000).
216. Suflita, M., Fu, L., He, W., Koffas, M. & Linhardt, R. J. Heparin and related polysaccharides: synthesis using recombinant enzymes and metabolic engineering. *Appl. Microbiol. Biotechnol.* **99,** 7465–7479 (2015).
217. Zhao, J. *et al.* Characterization of a novel modification of a CHO-produced mAb: Evidence for the presence of tyrosine sulfation. *MAbs* **9,** 985 (2017).
218. Garg, H. G. *et al.* Sulfation patterns in heparin and heparan sulfate: effects on the proliferation of bovine pulmonary artery smooth muscle cells. *Biochim. Biophys. Acta* **1639,** 225–231 (2003).
219. Swiech, K., Picanço-Castro, V. & Covas, D. T. Production of recombinant coagulation factors: Are humans the best host cells? *Bioengineered* **8,** 462 (2017).
220. Datta, P. *et al.* Bioengineered Chinese hamster ovary cells with Golgi-targeted 3-O-sulfotransferase-1 biosynthesize heparan sulfate with an antithrombin-binding site. *J. Biol. Chem.* **288,** 37308–37318 (2013).
221. Zhang, J. *et al.* High cell density cultivation of recombinant Escherichia coli strains expressing 2-O-sulfotransferase and C5-epimerase for the production of bioengineered heparin. *Appl. Biochem. Biotechnol.* **175,** 2986–2995 (2015).
222. He, W. *et al.* Production of chondroitin in metabolically engineered E. coli. *Metab. Eng.* **27,** 92–100 (2015).
223. Vázquez-Rey, M. & Lang, D. A. Aggregates in monoclonal antibody manufacturing processes. - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/21480193?dopt=Abstract. (Accessed: 22nd August 2018)
224. Paul, A. J., Handrick, R., Ebert, S. & Hesse, F. Identification of process conditions influencing protein aggregation in Chinese hamster ovary cell culture. *Biotechnol. Bioeng.* **115,** 1173–1185 (2018).
225. Jing, Y. *et al.* Identification of cell culture conditions to control protein aggregation of IgG fusion proteins expressed in Chinese hamster ovary cells. *Process Biochem.* **47,** 69–75 (2012).
226. Jenkins, N., Meleady, P., Tyther, R. & Murphy, L. Strategies for analysing and improving the expression and quality of recombinant proteins made in mammalian cells. *Biotechnol. Appl. Biochem.* **53,** 73–83 (2009).
227. Le Fourn, V., Girod, P.-A., Buceta, M., Regamey, A. & Mermod, N. CHO cell engineering to prevent polypeptide aggregation and improve therapeutic protein secretion. *Metab. Eng.* **21,** 91–102 (2014).
228. Tigges, M. & Fussenegger, M. Xbp1-based engineering of secretory capacity enhances the productivity of Chinese hamster ovary cells. - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/16635796. (Accessed: 22nd August 2018)
229. Borth N, E. al. Effect of increased expression of protein disulfide isomerase and heavy chain binding protein on antibody secretion in a recombinant CHO cell line. - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/15903247. (Accessed: 22nd

August 2018)
230. Mohan, C. & Lee, G. M. Effect of inducible co-overexpression of protein disulfide isomerase and endoplasmic reticulum oxidoreductase on the specific antibody productivity... - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/20506311. (Accessed: 22nd August 2018)
231. Chung JY, E. al. Effect of doxycycline-regulated calnexin and calreticulin expression on specific thrombopoietin productivity of recombinant Chinese hamster ovary c... - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/14760694. (Accessed: 22nd August 2018)
232. Hwang SO, E. al. Effect of doxycycline-regulated ERp57 expression on specific thrombopoietin productivity of recombinant CHO cells. - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/12573023. (Accessed: 22nd August 2018)
233. de Marco A, E. al. Native folding of aggregation-prone recombinant proteins in Escherichia coli by osmolytes, plasmid- or benzyl alcohol-overexpressed molecular chape... - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/16333986. (Accessed: 22nd August 2018)
234. Diamant S, E. al. Chemical chaperones regulate molecular chaperones in vitro and in cells under combined salt and heat stresses. - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/11517217. (Accessed: 22nd August 2018)
235. Schlapschy, M. & Skerra, A. Periplasmic Chaperones Used to Enhance Functional Secretion of Proteins in E. coli. in *Methods in Molecular Biology* 211–224 (2010).
236. Bierich, J. R. Treatment of pituitary dwarfism with biosynthetic growth hormone. *Acta Paediatr. Scand. Suppl.* **325,** 13–18 (1986).
237. Wadhwa, M. *et al.* Immunogenicity of granulocyte-macrophage colony-stimulating factor (GM-CSF) products in patients undergoing combination therapy with GM-CSF. *Clin. Cancer Res.* **5,** 1353–1361 (1999).
238. Beatson, R. *et al.* Transforming growth factor-β1 is constitutively secreted by Chinese hamster ovary cells and is functional in human cells. *Biotechnol. Bioeng.* **108,** 2759–2764 (2011).
239. Pavlovic, M., Girardin, E., Kapetanovic, L., Ho, K. & Trouvin, J.-H. Similar biological medicinal products containing recombinant human growth hormone: European regulation. *Horm. Res.* **69,** 14–21 (2008).
240. Zhang, Q. *et al.* Comprehensive tracking of host cell proteins during monoclonal antibody purifications using mass spectrometry. *MAbs* **6,** 659–670 (2014).
241. Levy, N. E., Valente, K. N., Choe, L. H., Lee, K. H. & Lenhoff, A. M. Identification and characterization of host cell protein product-associated impurities in monoclonal antibody bioprocessing. *Biotechnol. Bioeng.* **111,** 904–912 (2013).
242. Valente, K. N., Lenhoff, A. M. & Lee, K. H. Expression of difficult-to-remove host cell protein impurities during extended Chinese hamster ovary cell culture and their impact on continuous bioprocessing. *Biotechnol. Bioeng.* **112,** 1232–1242 (2015).
243. Levy, N. E., Valente, K. N., Lee, K. H. & Lenhoff, A. M. Host cell protein impurities in chromatographic polishing steps for monoclonal antibody purification. *Biotechnol. Bioeng.* **113,** 1260–1272 (2016).
244. Chiu, J. *et al.* Knockout of a difficult-to-remove CHO host cell protein, lipoprotein lipase, for improved polysorbate stability in monoclonal antibody formulations. *Biotechnol. Bioeng.* **114,** 1006 (2017).
245. Mamat, U. *et al.* Detoxifying Escherichia coli for endotoxin-free production of recombinant proteins. *Microb. Cell Fact.* **14,** (2015).
246. Chakrabarti, S., Barrow, C. J., Kanwar, R. K., Ramana, V. & Kanwar, J. R. Studies to Prevent Degradation of Recombinant Fc-Fusion Protein Expressed in Mammalian Cell Line and Protein Characterization. *Int. J. Mol. Sci.* **17,** (2016).
247. Hu, Z. *et al.* Carboxypeptidase D is the only enzyme responsible for antibody C-terminal lysine cleavage in Chinese hamster ovary (CHO) cells. *Biotechnol. Bioeng.* **113,** 2100–2106 (2016).
248. Overton, T. W. Recombinant protein production in bacterial hosts. *Drug Discov. Today* **19,** 590–601 (2014).
249. Jeong, H., Kim, H. J. & Lee, S. J. Complete Genome Sequence of Escherichia coli Strain BL21. *Genome Announc.* **3,** (2015).
250. Kerry-Williams, S. M., Gilbert, S. C., Evans, L. R. & Ballance, D. J. Disruption of theSaccharomyces cerevisiae YAP3 gene reduces the proteolytic degradation of secreted recombinant human albumin. *Yeast* **14,** 161–169 (1998).
251. Wu, M. *et al.* Disruption of YPS1 and PEP4 genes reduces proteolytic degradation of secreted HSA/PTH in Pichia pastoris GS115. *J. Ind. Microbiol. Biotechnol.* **40,** 589–599 (2013).
252. Doran, P. M. Foreign protein degradation and instability in plants and plant tissue cultures. *Trends Biotechnol.* **24,** 426–432 (2006).
253. Robert, S. *et al.* Protection of Recombinant Mammalian Antibodies from Development-Dependent Proteolysis in Leaves of Nicotiana benthamiana. *PLoS One* **8,** (2013).
254. Mandal, M. K., Fischer, R., Schillberg, S. & Schiermeyer, A. Inhibition of protease activity by antisense RNA improves recombinant protein production in Nicotiana tabacum cv. Bright Yellow 2 (BY-2) suspension cells. *Biotechnol. J.* **9,** 1065–1073 (2014).
255. Duwadi, K., Chen, L., Menassa, R. & Dhaubhadel, S. Identification, Characterization and Down-Regulation of Cysteine Protease Genes in Tobacco for Use in Recombinant Protein Production. *PLoS One* **10,** e0130556 (2015).
256. Gerngross, T. U. Corrigendum: Advances in the production of human therapeutic proteins in yeasts and filamentous fungi. *Nat. Biotechnol.* **22,** 1589–1589 (2004).
257. Gurramkonda, C. *et al.* Improving the recombinant human erythropoietin glycosylation using microsome supplementation in CHO cell-free system. *Biotechnol. Bioeng.* **115,** 1253–1264 (2018).
258. Jaroentomeechai T, E. al. Single-pot glycoprotein biosynthesis using a cell-free transcription-translation system enriched with glycosylation machinery. - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/30002445. (Accessed: 22nd August 2018)
259. Brödel, A. K., Sonnabend, A. & Kubick, S. Cell-free protein expression based on extracts from CHO cells. *Biotechnol. Bioeng.* **111,** 25–36 (2014).

CHAPTER 2

Exploring CRISPR/Cas9 multiplexing in CHO

Right before I started my PhD, our research group established a method for efficient CRISPR/Cas9 multiplexing of up to three CHO gene targets[1]. The simultaneous disruption of multiple target genes speeds up rational cell line development processes and is therefore more cost-effective than repetitive single knockout strategies. For the design of homogeneous N-glycan structures, often more than three genes are subject to engineering. Based on the above cited protocol we attempted the simultaneous disruption of ten gene targets in CHO cells. The results of the explorative study described in Chapter 2 were used as a guidance for upcoming cell line development strategies presented in Chapter 3 and Chapter 4.

1. Grav, L. M. *et al.* One-step generation of triple knockout CHO cell lines using CRISPR/Cas9 and fluorescent enrichment. *Biotechnol. J.* **10,** 1446–1456 (2015).

Deca CHO KO: exploring the limitations of CRISPR/Cas9 multiplexing in CHO cells

Thomas Amann[1], Anders Holmgaard Hansen[1], Gyun Min Lee[1,2], Mikael Rørdam Andersen[3], Helene Faustrup Kildegaard[1]

[1]The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark

[2]Department of Biological Sciences, KAIST, Daejeon, Republic of Korea

[3]Department of Biotechnology and Biomedicine, Technical University of Denmark, Kgs. Lyngby, Denmark

---

Abstract

CRISPR/Cas 9 multiplexing for simultaneous disruption of several target genes is a low-cost and time-saving method to generate genome-edited cell lines. Within N-glycosylation, there is a high number of possible gene targets and researches often have to disrupt all active isoforms within one protein class to remove certain enzymatic activities. Within CHO cells it was previously shown that multiplexing can be highly efficient for the disruption of up to three genes. In the presented study we performed multiplexing of 10 gene targets by simultaneous transfection of 10 sgRNA vectors and a GFP_2A_Cas9 encoding plasmid in a mAb-producing cell line. We then determined the efficiency of each sgRNA and sequenced the targeted genomic regions of 92 clonal cell lines. The sequence information revealed 20 clones with indels in one or more targets and one clone with indels in eight targeted genes. However the majority of gene-edited clones harboured a mixture of in-frame and out-of-frame shifts. Our study highlights that multiplexing several genes is a time saving, yet highly complex process challenged by transfection efficiency, sgRNA efficiency, and occurence of in-frame indels.

1 Introduction

As highlighted in chapter 1, the usage of genetic engineering tools in different expression platforms is steadily increasing and empowers the targeted design of superior host cell lines for the production of biopharmaceuticals with improved product qualities. Out of the available genetic engineering tools, CRISPR is very efficient for targeting single genes but also for simultaneous multiplexing of several targets during a single transfection. In 2015 Grav et al. reported the successful simultaneous generation of indels in three targeted genes in chinese hamster ovary (CHO) cells[1]. Within N-glycan engineering, researchers often aim to target more than three genes (typically in a sequential manner) in order to decrease N-glycan heterogeneity or to design very specific N-glycan structures[2]. The presence of numerous isoforms within glycosyltransferases calls for the disruption of more than one gene if a certain phenotype is desired. One example therefore is the class of β-1,4-galactosyltransferases (B4GALTs), consisting of B4GALT1–T7, which all transfer galactose from uridine diphosphate galactose (UDP-Gal) to N-Acetylglucosamine (GlcNAc) and GlcNAc-terminated oligosaccharides (EC 2.4.1.38)[3,4].

Step-wise disruption of several gene targets is often time consuming and therefore expensive. Within this study, the aim was to simultaneously disrupt ten CHO gene targets by CRISPR/Cas9 and thus to explore the limitations of the applied protocol. The ten selected target genes include genes involved in N-glycosylation (SPPL3, B4GALT1, B4GALT2, B4GALT3, B4GALT4, B4GALT5, TSTA3) and apoptosis/metabolism related genes (BAX, BAK, GLUL).

After co-transfection of a Cas9_2A_GFP encoding vector and ten sgRNAs encoding vectors, indel efficiencies for each sgRNA were analyzed on cell pool level before and after fluorescence-activated cell sorting (FACS) of GFP-positive cells. Additionally, 92 single cell clones from the same transfection were sequenced for disruptions in the targeted genes. In-frame and out-of-frame indel distribution of these clones as well as cell growth, mAb productivity and N-glycosylation of selected clones are presented.

2 Materials and methods

2.1 sgRNA and GFP_2A_Cas9 plasmid design

GFP_2A_Cas9 and single guide RNA (sgRNA) plasmids were constructed as previously described[1] and the sgRNA design was performed using CRISPy[5]. The target sites for BAX, BAK1, GLUL, SPPL3, B4GALT1, B4GALT2, B4GALT3, B4GALT4, B4GALT5, TSTA3 and the oligos for sgRNA cloning are

listed in Supporting Information, Table S1 and Table S2, respectively.

2.2 Cell cultivation, transfection and FACS for multiplexed genome editing

Rituximab producing CHO-S cells (generated by a protocol described earlier[6]) were cultivated, transfected for multiplexing and sorted for single cell clones via GFP-signal as previously described[7]. The GFP_2A_Cas9 / sgRNA plasmid ratios are presented in Table S3. To measure transfection efficiency, pmaxGFP® vector (Lonza, Basel, Switzerland) transfection was performed. Cells were harvested for fluorescence-activated cell sorting (FACS) 48 h after transfection.

2.3 Sampling for cell pool sequencing

Cell pools of transfections were sampled as pellets from ~200.000 cells by centrifugation (200 g, 5 minutes) 48 hours after transfection. A further sample of transfected cell pools was taken 48 hours after transfection by generating a cell pool (~200.000 cells) of GFP-positive cells via FACS.

2.4 Deep sequencing analysis

Cell pool samples and confluent colonies from 96-well flat-bottom replicate plates were prepared for sequencing and analyzed on a MiSeq Benchtop Sequencer (Illumina, San Diego, CA) as described previously[1]. PCR primers are presented in Supporting information, Table S4.

2.5 Batch cultivation to study cell growth and secretome N-glycan analysis

Cells were seeded in duplicates at 3.0 x 10^5 cells/mL in Corning vent cap shake flasks (Sigma-Aldrich, St. Louis, MI) in 30 mL CD CHO medium supplemented with 8 mM L-glutamine and 1 µL/mL anti-clumping agent (Life Technologies). Cells were incubated for seven days in a humidified incubator at 120 rpm, 37°C and 5% CO_2. Cell densities and cell viabilities were determined daily using the NucleoCounter NC-250 Cell Counter (ChemoMetec). Additionally rituximab titers were assessed daily and a 5 mL N-glycan sample was pooled within duplicates at day 4 for secretome (total secreted host cell protein) analysis.

2.6 Rituximab quantification and secretome N-glycan analysis

Secretome samples were analyzed for N-glycan structures via a LC-MS as previously described[7]. Rituximab quantification was performed with an Octet RED96 (Pall, Menlo Park, CA) as previously described[8].

2.7 Batch cultivation to study cell growth in L-gln-free medium

Clones were seeded in duplicates at 3 x 10^5 cells/mL in 96 half-deepwell plates (CR1496c, Enzyscreen, Haarlem, the Netherlands). Cell cultivation and daily determination of cell viability and cell density was performed as previously described[6,9]. For cultivation, CD CHO medium was supplemented with 8 mM L-glutamine and 1 µL/mL anti-clumping agent (Life Technologies) or only supplemented with 1 µL/mL anti-clumping agent.

3 Results

3.1 Indel efficiencies of sgRNAs within sorted and unsorted cell pools

48 hours after the transfection of GFP_2A_Cas9 and ten sgRNAs GFP-positive cells were bulk sorted (~32% of total population). Afterwards, samples of unsorted and GFP-sorted cell pools were analyzed for insertions or deletions (indels) in the targeted sequences. The comparison of indel efficiencies for each target before and after sorting are presented in Figure 1. Indel efficiency in unsorted cells ranged from 0.8% (GLUL) to 38.2% (TSTA3) and was elevated to a range from 1.6% (GLUL) to 80.1% (TSTA3) after FACS. By sorting for GFP-positive cells, the frequency of indel generation increased at least one-fold for all sgRNAs. Compared to the indel efficiencies in other targets, GLUL indel percentage was relatively low whereas sgRNAs targeting BAX, BAK, TSTA3, B4GALT2, B4GALT3 and B4GALT4 lead to indels in >50% of the sorted cells.

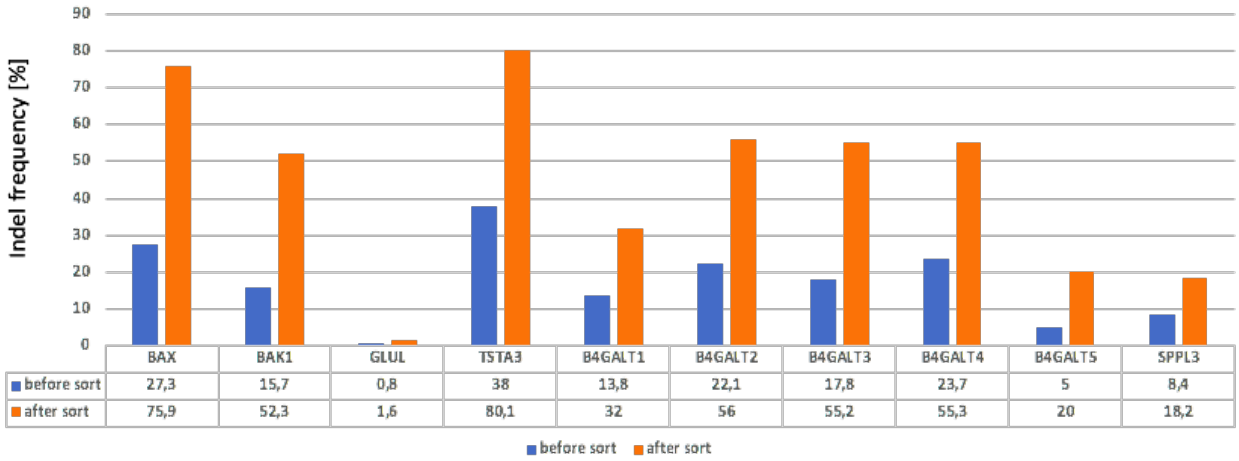


Figure 1. Indel efficiency within cell pools. The sgRNA indel efficiency for each target before (blue) and after (orange) FACS is presented. Values represent the percentage of identified indel sequences within the total cell population.

3.2 Disruption of targeted genes in single cell clones

Besides analyzing cell pools, single cell clones were also investigated for gene disruptions. After transfection and FACS-based single cell cloning, the target sites of 92 clones were sequenced, whereof 20 clones were found to harbour indels in one or more targeted genes (Figure 2 and Supporting Information, Table S5). No clone was identified to have an indel in GLUL. However four clones that were not able to grow in L-glutamine-free medium were identified, indicating a successful disruption of GLUL. The amount of disrupted target genes per clone ranged from one to eight (clone 68). A total of 53 indels were found in the 20 clones, thereof 34 were out-of-frame (blue) and 19 in-frame (grey) indels. For BAX, BAK, B4GALT3 and B4GALT4 more out-of-frame than in-frame disruptions were identified. Thereof the sgRNA against BAX generated the highest amount of out-of-frame indels. For B4GALT1 there were no indels found. sgRNAs against TSTA3, B4GALT2, B4GALT5 and SPPL3 generated more in-frame than out-of-frame indels or no out-of-frame indels at all. We selected three clones for further characterization: clone 36 with indels in four targeted genes, clone 68 in eight targeted genes plus the GLUL disruption, and clone 90 with indels in six targeted genes (asterisk clones in Figure 2).

	BAX	BAK	GLUL	TSTA3	B4GALT1	B4GALT2	B4GALT3	B4GALT4	B4GALT5	SPPL3	total indels		
clone nr			Phenotype								out-of frame	in-frame	total
5											1	0	1
7											0	1	1
14											2	0	2
16											3	1	4
23											1	0	1
36*											3	1	4
40											1	3	4
49											1	2	3
57											2	0	2
60											1	0	1
68*											3	5	8
73											1	0	1
76											2	0	2
77											2	0	2
78											2	1	3
88											3	1	4
90*											3	3	6
91											0	1	1
92											2	0	2
93											1	0	1
out-of frame	13	7	4	2	0	0	2	10	0	0	**34**	**19**	**53**
in-frame	3	2	-	8	0	2	1	1	1	1			

Figure 2. Analysis of target gene disruptions. MiSeq analysis identified out-of-frame (blue) and in-frame (grey) indels in the targeted genes. Assumed gene disruption of GLUL (no growth in L-Gln-free medium) is indicated in green. Clones that were selected for further analysis are marked by an asterisk.

3.3 Growth and productivity of single cell clones

To investigate whether the multiplexing interfered with cell growth or mAb productivity, selected clones with indels, two control clones without any indels (CTR1, CTR2) and the parental cell line were cultivated in batch format for seven days. As displayed in Figure 3A, maximum viable cell densities (VCD) ranged from 2 x 10^6 cells/mL (CTR1, CTR2, clone 68) to 6 x 10^6 cells/mL (clone 36). With exception of clone 36 and the parental cell line, all clones maintained their maximum VCD two to four days.

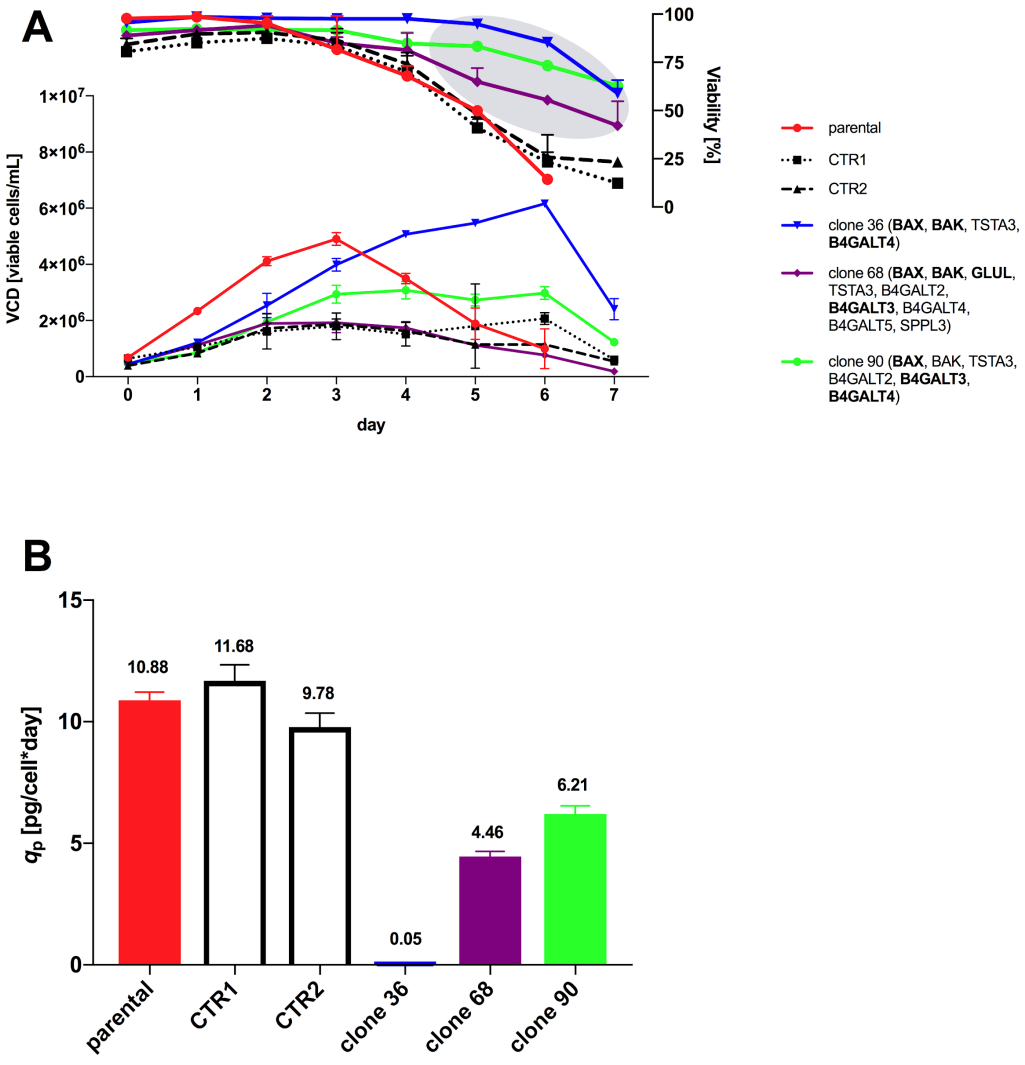


Figure 3. Cell growth and mAb productivity during batch cultivation. (A) Daily viable cell densities and cell viabilities. The grey area displays the slow decline in viabilities for clones with indels in BAX and BAK1. Targets with identified indels are shown in brackets. Bold target names represent out-of-frame indels. (B) Specific mAb productivities q_p. Rituximab productivity was determined for each cell line during the exponential growth phase and is displayed as pg/cell*day. Error bars indicate range of shake flask duplicates.

For the parental cell line, CTR1 and CTR2 the cell viabilities declined rapidly after four days of cultivation. In contrast, the clones with indels in BAX and BAK (clone 36, 68 and 90) revealed a slow decline in cell viabilities between day four and seven (grey area, Fig. 3A).

After daily quantification of mAb in the culture supernatants, the specific productivity (q_p) was calculated for each clone during the exponential growth phase (Figure 3B). The assessed q_p for both control clones was similar to the parental clone (10.88 pg/cell*day) whereas the three clones with gene target indels had decreased q_p (clone 68: 4.46 pg/cell*day, clone 90: 6.21 pg/cell*day) or lost mAb productivity completely (clone 36: 0.05 pg/cell*day).

3.4 N-glycan analysis of single cell clones

In order to assess whether the generated indels in the N-glycosylation-related target genes had an effect on the N-glycan profile of the cell lines and if the protocol for cell line development (CLD) influences overall N-glycosylation, supernatant samples harboring total secreted protein were analyzed for N-glycan structures.

It was shown that the secretome N-glycans of control clones were similar to the counterpart of the parental cell line (Supporting Information, Figure S1). The three clones with indels revealed slightly less N-glycan heterogeneity, but similar core-fucosylation to the control cell lines as exemplified in the direct comparison of annotated N-glycans from CTR1 and clone 68 with indels in BAX, BAK, SPPL3, B4GALT2/3/4/5 and TSTA3 (Figure 4). The dominant N-glycan found in all cell lines was the fully capped and core-fucosylated bi-antennary A2FS2 structure, followed by the A2FS1 N-glycan.

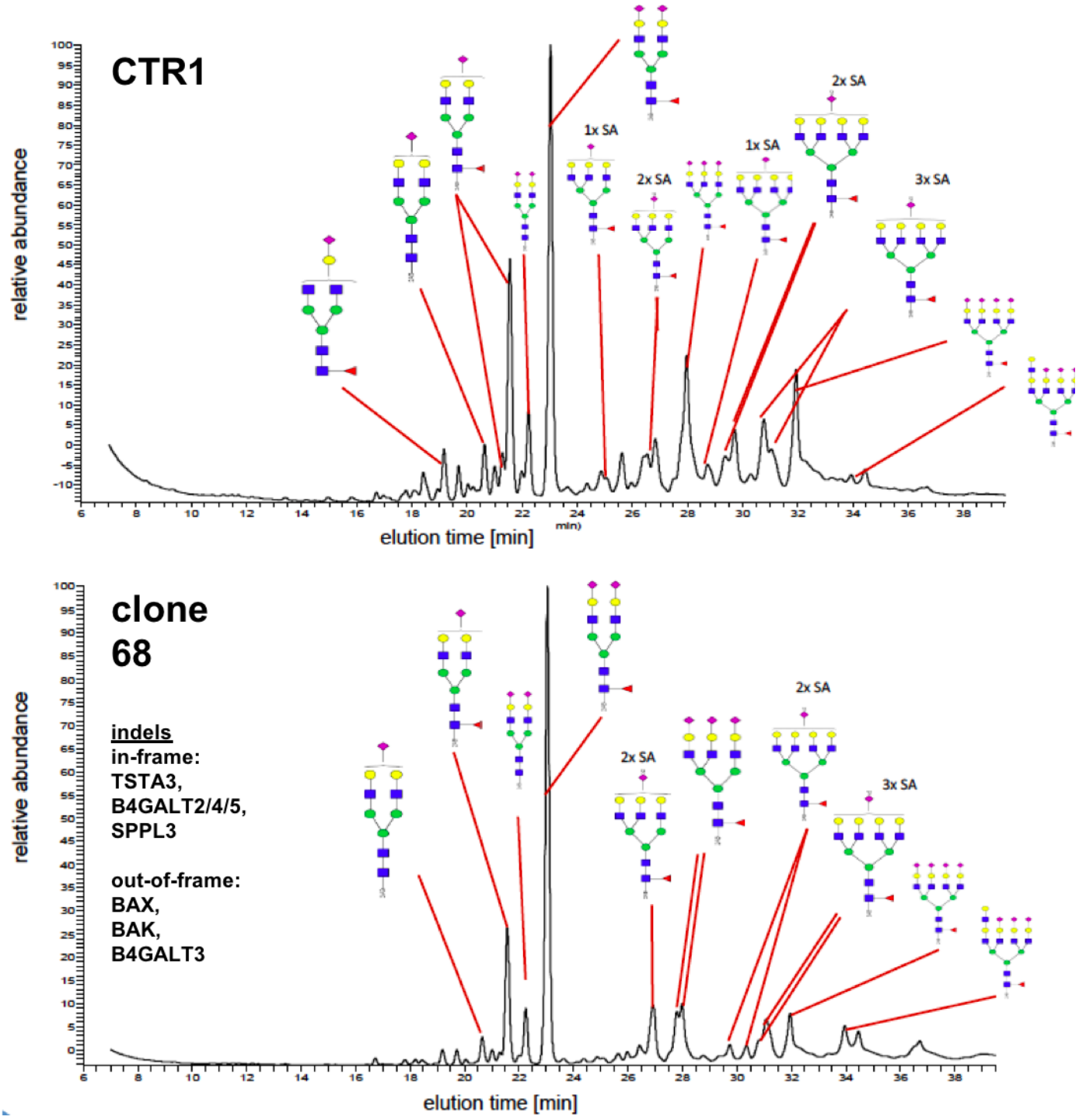


Figure 4. N-glycan analysis of secretome. Released N-glycans from total secreted proteins were analysed and annotated for CTR1 (no indels) and clone 68 (indels in 8 targeted genes) (SA=sialic acid).

4 Discussion

In this study we present CRISPR/Cas9 multiplexing of ten gene targets in CHO cells. The motivation for the study was to save time and therefore resources during CLD towards improved cell lines for biopharmaceutical production. We explored the feasibility of simultaneous transfection of 11 DNA vectors (10 sgRNA and one Cas9-encoding plasmid), high-throughput sequencing of 10 target sequences in 92 clones and following characterization of multi-indel clones.

The average ratio between *in-frame : out-of-frame* indels of the targeted sequences from all 92 clones was as expected close to 1 : 2 (19 : 34). Therefore, by planning multiplexing several gene targets, the theoretical amount of clones to screen in order to identify one clone with pure out-of-frame indels can be calculated with equation (1) of Figure 5.

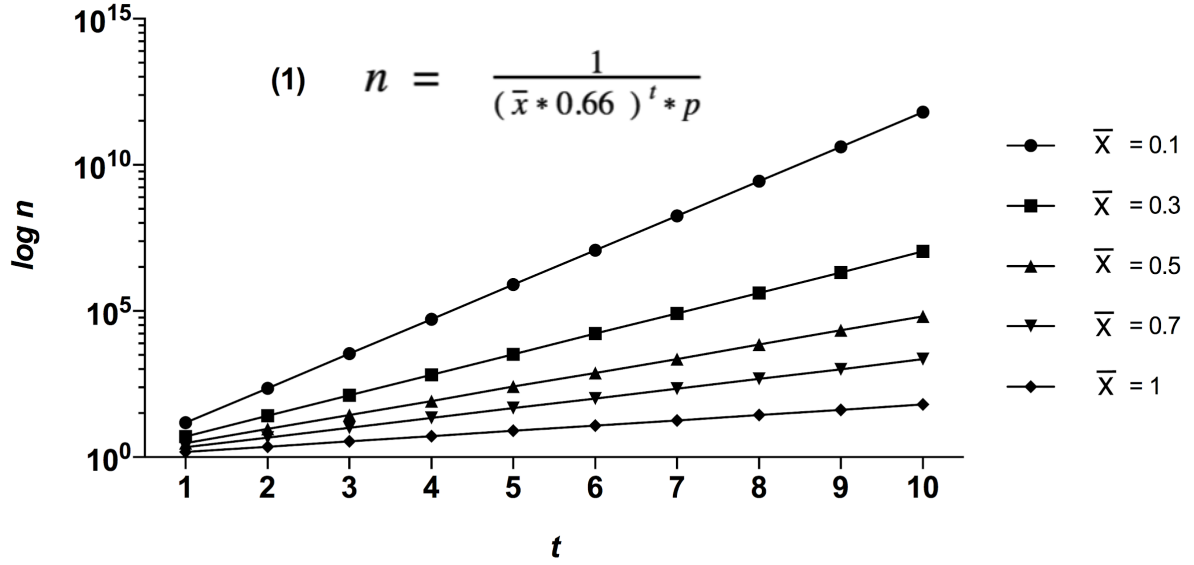


Figure 5. Theoretical determination of clone screening volume via equation (1). The amount of clones to screen to identify one clone with exclusively out-of-frame indels (*n*) is dependent on the average indel frequency ($\bar{x}$), number of targeted genes (*t*) and uptake rate of all genetic elements by the transfected cell (*p*). For above illustration the transfection efficacy was set to *p*=1.

Figure 5 highlights the importance of indel frequency and number of chosen target genes, where equation (1) enables an estimation of the amount of clones to screen to identify a clone with out-of-frame disruptions in all targeted genes. Additionally, the increasing plasmid number for transfection during the multiplexing is negatively correlated to the probability (*p*) of cells that successfully took up all transfected plasmids.

As presented in section 3.1, the event of indel generation can be boosted by sorting for transfected cells expressing GFP_2A_Cas9. Using the indel frequencies of the GFP sorted cell pools in this study ($\bar{x}$=0.446) and p = 1, a number of 2.04 x 10^5 cells needs to be sequenced to hypothetically identify one clone with exclusively out-of-frame indels in all ten targeted genes. Without GFP-sorting ($\bar{x}$=0.173) this number increases to 2.66 x 10^9 cells. If, again hypothetically, the indel frequency could be increased to 100% for each sgRNA ($\bar{x}$=1.0) the amount of clones to screen to identify a 10x KO clone with exclusively out-of-frame indels would drop to only 64 (Figure 5).

However 100% indel frequencies are not realistic yet. To successfully generate cell lines with out-of frame indels in up to ten targets we suggest a step-wise approach. For instance, one possible strategy is 3 rounds of multiplexing, where 3 - 5 genes are targeted in each round. The number of clones to screen would then decrease from 2.04 x 10^5 (10 targets) to 455 (5 targets) or 30 (3 targets) per round.
Sakuma et al. reported that efficient indel generation during multiplexing seven gene targets can be enhanced by encoding all sgRNAs with the Cas9 protein on a single vector[10]. However the study did not distinguish between in-frame and out-of-frame indels or generate and characterize single cell clones. As published recently, easily customizable, reliable and cost-effective protocols for the generation of vectors with high numbers of sgRNAs are available for multiplexing approaches[11,12]. Also, the frequency of indel generation in CHO can be doubled by increasing the expression levels of both, Cas9 and sgRNA[13].
The importance of correct sequence annotation and information for the optimal design of sgRNAs is highlighted in this study by the example of targeting the GLUL gene. Although we could not confirm the presence of indels in the targeted GLUL sequence, we identified clones with L-Gln auxotrophy. Therefore we suggest that the designed sgRNA generated a disruption in an unannotated but expressed GLUL gene variant, whereas the originally targeted GLUL sequence might actually be a pseudogene.

In our study, the engineered clones were found to have decreased N-glycan heterogeneity and lower mAb productivity than the parental cell line. Also, the cell growth of all engineered clones was different to the parental cell line. However it is challenging to determine whether a disrupted gene sequence with in-frame indel leads to a non-functional protein - and therefore to a certain phenotype - or not.
The two control clones without indels in the targeted genes had lower max. VCD than the parental cell line, but did not reveal altered N-glycosylation or q_p. Unlike N-glycosylation, we suggest that cell growth can be affected by the presented CLD protocol. To not compromise q_p due to possibly unstable

transgene integration, cell engineering by multiplexing should be performed prior integrating the gene of interest and not vice versa.

We conclude that simultaneous CRISPR/Cas9 multiplexing of a high number of CHO target genes can be a time saving method. However the amount of clones to screen increases exponentially with the number of targeted genes. The generation of undesired in-frame indels decreases overall efficiency, whereas improved sgRNA (plasmid) design could keep screening efforts in a practical range.

Abbreviations:

B4GALT, β-1, 4-Galactosyltransferase; **BAK**, BCL2 Antagonist/Killer1; **BAX**, BCL2 Associated X; **Cas9**, CRISPR associated protein 9; **CHO**, chinese hamster ovary; **CLD**, cell line development; **CRISPR**, clustered regularly interspaced short palindromic repeats; **FACS**, fluorescence-activated cell sorting; **GFP**, green fluorescent protein; **GlcNAc**, N-Acetylglucosamine; **GLUL**, glutamate-ammonia ligase; **indel**, gene insertion/deletion; **LC-MS**, liquid chromatography-mass spectrometry; **L-Gln**, L-glutamine; **mAb**, monoclonal antibody; **qp**, cell specific productivity; **sgRNA**, single guide RNA; **SPPL3**, signal peptide peptidase like 3; **TSTA3**, tissue specific transplantation antigen P35B; **VCD**, viable cell density

5 References

1. Grav, L. M. *et al.* One-step generation of triple knockout CHO cell lines using CRISPR/Cas9 and fluorescent enrichment. *Biotechnol. J.* **10,** 1446–1456 (2015).
2. Yang, Z. *et al.* Engineered CHO cells for production of diverse, homogeneous glycoproteins. *Nat. Biotechnol.* **33,** 842–844 (2015).
3. Amado, M., Almeida, R., Schwientek, T. & Clausen, H. Identification and characterization of large galactosyltransferase gene families: galactosyltransferases for all functions. *Biochim. Biophys. Acta* **1473,** 35–53 (1999).
4. Furukawa, K. & Sato, T. Beta-1,4-galactosylation of N-glycans is a complex process. *Biochim. Biophys. Acta* **1473,** 54–66 (1999).
5. Ronda, C. *et al.* Accelerating genome editing in CHO cells using CRISPR Cas9 and CRISPy, a web-based target finding tool. *Biotechnol. Bioeng.* **111,** 1604–1616 (2014).
6. Pristovšek, N. *et al.* Using Titer and Titer Normalized to Confluence Are Complementary Strategies for Obtaining Chinese Hamster Ovary Cell Lines with High Volumetric Productivity of Etanercept. *Biotechnol. J.* **13,** e1700216 (2018).
7. Amann, T. *et al.* CRISPR/Cas9-Multiplexed Editing of Chinese Hamster Ovary B4Gal-T1, 2, 3, and 4 Tailors N-Glycan Profiles of Therapeutics and Secreted Host Cell Proteins. *Biotechnol. J.* (2018). doi:10.1002/biot.201800111
8. Kallehauge, T. B. *et al.* Ribosome profiling-guided depletion of an mRNA increases cell growth rate and protein secretion. *Sci. Rep.* **7,** (2017).
9. Hansen, H. G. *et al.* Versatile microscale screening platform for improving recombinant protein productivity in Chinese hamster ovary cells. *Sci. Rep.* **5,** 18016 (2015).
10. Sakuma, T., Nishikawa, A., Kume, S., Chayama, K. & Yamamoto, T. Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. *Sci. Rep.* **4,** 5400 (2014).
11. Breunig, C. T. *et al.* One step generation of customizable gRNA vectors for multiplex CRISPR approaches through string assembly gRNA cloning (STAgR). *PLoS One* **13,** e0196015 (2018).
12. Kurata, M. *et al.* Highly Multiplexed Genome Engineering Using CRISPR/Cas9 gRNA Arrays. (2018). doi:10.1101/331371
13. Shin, J. *et al.* Efficient CRISPR/Cas9-mediated multiplex genome editing in CHO cells via high-level sgRNA-Cas9 complex. *Biotechnol. Bioprocess Eng.* **20,** 825–833 (2015).

CHAPTER 3

Decreasing CHO N-glycan galactosylation

A major advantage of expressing therapeutic proteins in CHO cells is their capability to produce complex type N-glycosylation similar to humans. However, the N-glycan structures of CHO cells are highly heterogeneous leading to inconsistent product quality. The disruption of N-glycosyltransferases can decrease this heterogeneity and researchers can design certain N-glycan profiles which potentially have beneficial effects on efficacy or pharmacokinetics of therapeutic proteins. In this chapter we describe the disruption of beta-1,4-galactosyltransferase isoforms by CRISPR/Cas9 multiplexing to produce non-galactosylated glycoproteins and to elucidate the contribution of each isoform to CHO-S N-glycan galactosylation.

RESEARCH ARTICLE

CRISPR/Cas9

Biotechnology Journal

www.biotechnology-journal.com

CRISPR/Cas9-Multiplexed Editing of Chinese Hamster Ovary B4Gal-T1, 2, 3, and 4 Tailors *N*-Glycan Profiles of Therapeutics and Secreted Host Cell Proteins

Thomas Amann, Anders Holmgaard Hansen, Stefan Kol, Gyun Min Lee, Mikael Rørdam Andersen, and Helene Faustrup Kildegaard**

In production of recombinant proteins for biopharmaceuticals, *N*-glycosylation is often important for protein efficacy and patient safety. IgG with agalactosylated (G0)-*N*-glycans can improve the activation of the lectin-binding complement system and be advantageous in the therapy of lupus and virus diseases. In this study, the authors aimed to engineer CHO-S cells for the production of proteins with G0-N-glycans by targeting B4Gal-T isoform genes with CRISPR/Cas9. Indel mutations in genes encoding B4Gal-T1, -T2, and -T3 with and without a disrupted B4Gal-T4 sequence resulted in only ≈1% galactosylated *N*-glycans on total secreted proteins of 3-4 clones per genotype. The authors revealed that B4Gal-T4 is not active in *N*-glycan galactosylation in CHO-S cells. In the triple-KO clones, transiently expressed erythropoietin (EPO) and rituximab harbored only ≈6% and ≈3% galactosylated *N*-glycans, respectively. However, simultaneous disruption of B4Gal-T1 and -T3 may decrease cell growth. Altogether, the authors present the advantage of analyzing total secreted protein *N*-glycans after disrupting galactosyltransferases, followed by expressing recombinant proteins in selected clones with desired *N*-glycan profiles at a later stage. Furthermore, the authors provide a cell platform that prevalently glycosylates proteins with G0-*N*-glycans to further study the impact of agalactosylation on different in vitro and in vivo functions of recombinant proteins.

1. Introduction

Chinese hamster ovary (CHO)-derived cells are the major workhorses within mammalian cell lines and represent the cell platform in which >50% of the marketed recombinant proteins are produced.[1] Thereof, recombinant monoclonal antibodies (mAbs) are the main product subclass and are utilized for the treatment of cancer and various inflammatory diseases.[2] As a result of post-translational protein processing, mAbs harbor two predominantly bi-antennary *N*-glycans, one on each heavy chain at Asparagine (Asn) 297. In contrast to mAbs, erythropoietin (EPO) has three *N*-glycosylation sites occupied by predominantly tri- and tetra-antennary structures.[3] In general, *N*-glycosylation can impact protein folding, immune regulation, cellular homeostasis and the biological half-life of proteins.[4,5] Within mAbs, the fragment crystallizable (Fc) *N*-glycans at Asn297 have a strong influence on anti-inflammatory properties, antibody-dependent cell-mediated cytotoxicity and complement-dependent cytotoxicity.[6]

The heterogeneous *N*-glycan profile of glycoproteins produced in CHO is one of the main factors that causes mAb heterogeneity and can be further optimized regarding core-fucosylation, galactosylation, antennarity, and terminal capping by sialic acids. Rituximab is an immunoglobulin G (IgG) 1-class molecule, one of the recombinant glycoproteins produced in CHO, and exceeds annual revenues of USD 7 billion.[7] Rituximab targets the B-cell surface antigen CD20 in B-cell lymphoma and is predominantly *N*-glycosylated by A2FG0 and A2FG1 structures when produced in non-glycoengineered CHO cells.[8] Since several studies revealed nonfucosylated IgGs have significantly higher binding affinity for the Fc-gamma receptor IIIa (FcγRIIIa) than fucosylated IgG versions,[9,10] different approaches successfully removed the core-fucose by knockout of alpha-(1,6)-fucosyltransferase (FUT8) or tissue-specific transplantation antigen P35B (TSTA3) in IgG-expressing CHO cell lines.[11–14]

Additionally, agalactosylated IgG1 variants with terminal *N*-Acetylglucosamine (GlcNAc) (referred to as G0 glycoforms)

T. Amann, A. H. Hansen, Dr. S. Kol, Prof. G. M. Lee, Dr. H. F. Kildegaard
The Novo Nordisk Foundation Center for Biosustainability
Technical University of Denmark
Kemitorvet, Building 220, 2800 Kgs. Lyngby, Denmark
E-mail: ahoha@biosustain.dtu.dk; hef@biosustain.dtu.dk

Prof. G. M. Lee
Department of Biological Sciences
KAIST
Daejeon, Republic of Korea

Prof. M. R. Andersen
Department of Biotechnology and Biomedicine
Technical University of Denmark
Kgs. Lyngby, Denmark

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/biot.201800111.

DOI: 10.1002/biot.201800111

can increase the binding to FcRIIIa[15] and are accessible for the mannose-binding protein. They can therefore promote activation of the lectin-binding complement system[16] without impacting in vivo clearance.[17–19] Furthermore, HIV patients with high viral inhibition displayed an increased proportion of agalactosylated *N*-glycans on global serum IgG, suggesting that agalactosylated IgG variants may have antiviral activity.[20] Interestingly, lupus patients showed improved disease symptoms after treatment with agalactosylated antibodies.[21] These G0-IgG variants can be obtained by sequential treatment of wild type (WT)-IgG with neuraminidase and galactosidase or by supplementing the cultivation medium with galactose analogues to block cellular B4Gal-Ts.[22] Nevertheless, fewer cell engineering attempts were initiated to produce G0-IgG1 compared to engineering nonfucosylated IgG1 variants.

As the CHO genome sequence became publicly available,[23] CHO cell engineering is no longer performed in a "black box," which shortens cell line development and empowers a targeted approach for the engineering of a G0 CHO cell line. The classes of glycosyltransferases are made of homologous gene families, where the class of β-1,4-galactosyltransferases (B4GalTs) consists of seven members, B4Gal-T1–T7, which all transfer galactose from uridine diphosphate galactose (UDP-Gal) to GlcNAc and GlcNAc-terminated oligosaccharides (EC 2.4.1.38).[24,25] B4Gal-T5 and -T6 are described to mainly function in O-glycosylation,[26,27] whereas B4Gal-T7 transfers UDP-Gal within glycosaminoglycan biosynthesis and, therefore, is not involved in the *N*-glycosylation of proteins.[28,29] A further study indicated that B4Gal-T1, -T2, -T3, and -T4 perform *N*-glycan galactosylation more efficiently than B4Gal-T5 and -T6 and suggested different branch preferences for the family members of β-1,4-galactosyltransferases.[30] In addition, B4Gal-T4 is reported to also be active in the galactosylation of mucin-type core 2 branching in the O-glycosylation pathway.[31] Furthermore, B4Gal-T1-KO mutants are described to have dramatically reduced galactosylation on secreted host cell proteins (secretome) *N*-glycans and reduced growth of mice.[27,32] In a previous study performed with CHO-K1 cells, disruption of B4Gal-T1, -T2, and -T3 led to almost fully agalactosylated EPO and rituximab.[33] However, the impact of B4Gal-T disruptions on cell growth of more than one clone was not performed. The CHO-K1 study included single-KO of the B4Gal-T isoforms targeted in our work. In contrast, we aimed to study the *N*-glycosylation activity of B4Gal-T1, -T2, -T3, and -T4 after combinatorial KO in the industrially relevant CHO-S cell line. Especially, the role of B4Gal-T2 and -T4 in CHO-S and the effect of B4Gal-T indels on cell growth, both with respect to clonal variation, were the driving motives of this work. Therefore, we applied clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9 (CRISPR/Cas9) to multiplex B4Gal-T disruptions in CHO-S cells. Cell growth and protein *N*-glycosylation profiles of multiple clones for each triple- and quadruple-KO combinations were analyzed to additionally examine clonal variation. *N*-glycosylation analysis of total secreted proteins, as well as transiently expressed rituximab and EPO (representing dissimilar *N*-glycan profiles), in the B4Gal-T edited cell lines was performed. The analysis demonstrated that *N*-glycans can be tailored for a greater variety of secreted glycoproteins, as represented by more than 250 proteins within the CHO-S secretome[34] in addition to EPO and rituximab. With this, we found that screening the secretome *N*-glycans of our engineered clones is a promising strategy toward the expression of rituximab and EPO with G0 *N*-glycans in selected clones.

2. Experimental Section

2.1. sgRNA and GFP_2A_Cas9 Plasmid Design

GFP_2A_Cas9 and single-guide RNA (sgRNA) plasmids were constructed as previously described.[13] The sgRNA target design for B4Gal-T1, B4Gal-T2, B4Gal-T3, and B4Gal-T4 was performed using CRISPy.[35] The target sites for the mentioned genes and the oligos for sgRNA cloning are listed in Tables S1 and S2, Supporting Information, respectively.

2.2. Cell Cultivation and Transfection for Multiplexed Genome Editing

CHO-S suspension cells (Life Technologies, Carlsbad, CA) were cultivated in a CD CHO medium supplemented with 8 mM L-glutamine and 1 μL mL^{-1} anticlumping agent (Life Technologies). Cells were incubated in a humidified incubator at 120 rpm, 37 °C and 5% CO_2. Cell passaging was conducted every 2–3 days at 3×10^5 cells mL^{-1} after measuring viable cell densities (VCDs) and viabilities with the NucleoCounter NC-200 Cell Counter (ChemoMetec, Allerod, Denmark). One day prior transfection with CRISPR reagents, the anticlumping agent was removed by centrifugation and 5 to 6×10^5 cells mL^{-1} were seeded in a six well plate (BD Biosciences, San Jose, CA) for each transfection. At the day of transfection, each sample was seeded at 1×10^6 cells mL^{-1}, and a total DNA load of 3.5 μg was transfected with FuGENE® HD transfection reagent (Promega, Madison, WI) and OptiPRO SFM medium (Life Technologies), according to the manufacturer's recommendations. The GFP_2A_Cas9/sgRNA plasmid ratios for each sample are presented in Table S3, Supporting Information. To measure transfection efficiency, pmaxGFP® vector (Lonza, Basel, Switzerland) transfection was performed. Cells were harvested for fluorescence-activated cell sorting (FACS) 48 h after transfection.

2.3. Single-Cell Cloning Using FACS

Before FACS, cells were filtered through a 40 μm cell strainer into a FACS-compatible tube.

Operating a FACSJazz (BD Biosciences), single fluorescent-positive cells were sorted into 384-well plates (Corning, New York, NY) already containing 30 μL CD CHO medium supplemented with 8 mM L-glutamine, 1.5% HEPES buffer and 1% Antibiotic-Antimycotic (Gibco, Waltham, MA) per well. For cell sorting, fluorescent-positive cell populations were gated based on nontransfected WT CHO-S cells. Two weeks after cell sorting the clones were moved to 96-well flat-bottom plates (BD Biosciences) and expanded for deep sequencing analysis and batch cultivation.

2.4. Deep Sequencing Analysis

Confluent colonies from 96-well flat-bottom replicate plates were harvested for genomic DNA extraction. DNA extraction was performed using QuickExtract DNA extraction solution (Epicentre, Illumina, Madison, WI) according to the manufacturer's instructions. The library preparation was based on Illumina 16S Metagenomic Sequencing Library Preparation, and deep sequencing was carried out on a MiSeq Benchtop Sequencer (Illumina, San Diego, CA). The protocol for amplifying the targeted genomic sequences, amplicon purification, adapter-PCR and following quality analysis was based on a previously published work.[13] PCR primers are presented in Table S4, Supporting Information.

2.5. Batch Cultivation to Study Cell Growth and Secretome *N*-Glycans

For batch cultivation and secretome analysis, cells were seeded at 3.0×10^5 cells mL^{-1} in Corning vent cap shake flasks (Sigma–Aldrich, St. Louis, MI) as duplicates in 30 mL CD CHO medium supplemented with 8 mM L-glutamine and 1 $\mu L\,mL^{-1}$ anti-clumping agent (Life Technologies). Cells were incubated in a humidified incubator at 120 rpm, 37 °C and 5% CO_2. Cell densities and viabilities were determined once per day using the NucleoCounter NC-250 Cell Counter (ChemoMetec). Based on cell densities from days 0 to 7 we calculated the integral of viable cells (IVC). The IVC was used to statistically determine differences in cell growth between the generated clones. Using a two-tailed, unpaired *t*-test with Prism7 software, we grouped the clones into two sets. One set with ($n = 16$) and one set without combinatorial disruption of B4Gal-T1 and -T3 ($n = 16$). Combining all IVC values of the clones in each set, we determined if a set of clones had a significant change of IVC compared to CHO-S WT and WT ctr cells. Secretome sample volume was calculated to harbor 20×10^6 cells and harvested 5 days after seeding to be pooled within biological replicates.

2.6. Batch Cultivation for Transient Rituximab/EPO Transfection and Rituximab/EPO *N*-Glycan Analysis

For transient expression of rituximab and EPO, cells were seeded in Corning vent cap shake flasks (Sigma–Aldrich) as duplicates with cell densities $\approx 1 \times 10^6$ cells mL^{-1} in 60 mL CD CHO medium supplemented with 8 mM L-glutamine (Life Technologies). Cells were incubated in a humidified incubator at 120 rpm, 37 °C and 5% CO_2 and transfected with 75 μg of rituximab or EPO encoding plasmid for each flask using FreeStyle™ MAX reagent together with OptiPRO SFM medium (Life Technologies) according to the manufacturer's recommendations. A total of 1 $\mu L\,mL^{-1}$ anticlumping agent was added 24 h after transfection. pmaxGFP® vector (Lonza) transfection was performed to measure transfection efficiencies. Cell densities and viabilities were determined once per day using the NucleoCounter NC-250 Cell Counter (ChemoMetec). To purify rituximab and EPO, the supernatants of the transfected clones were harvested 3 days after transfection and pooled within duplicates.

2.7. Rituximab and EPO Purification

For rituximab purification, supernatant samples were centrifuged (1000 g, 5 min, 4°C) and afterwards filtered (≈0.22 μm pore size) to remove cells and cell debris. Rituximab was purified by protein A affinity chromatography (MabSelect, GE Healthcare, Uppsala, Sweden) according to the manufacturer's protocol. Human protein C4 (HPC4)-tagged EPO was purified from supernatants using Anti-Protein C Affinity Matrix from Roche (Basel, Switzerland, Cat. Nr. 11815024001) as per the instructions of the manufacturer.

2.8. *N*-Glycan Analysis

Sample preparation for *N*-glycan analysis was performed with GlycoWorks RapiFluor-MS *N*-Glycan Kit (Waters, Milford, MA) according to the manufacturer's instructions. A total of 12 μg purified protein or 12 μL of 10× concentrated (Amicon Ultra-15, Merck, Darmstadt, Germany) secretome sample were used for each sample. Labeled *N*-Glycans were analyzed by a LC-MS system using a Thermo Ultimate 3000 HPLC with fluorescence detector coupled online to a Thermo Velos Pro Iontrap MS, as described previously with minor modifications.[13] Separation gradient was 30% to 43% buffer, and MS was run in positive mode. The amount of *N*-Glycan was measured by integrating the areas under the normalized fluorescence spectrum peaks with Thermo Xcalibur software (Thermo Fisher Scientific, Waltham, MA) giving the normalized, relative amount of the glycans.

3. Results

3.1. Generation of Engineered CHO-S Cell Lines With Combinations of Indels in Multiple B4Gal-T Genes

To investigate the exact impact of B4Gal-T1, -T2, -T3, and -T4-KO on *N*-glycan galactosylation, we aimed to generate clones with insertion or deletion (indel) mutations in several of the genes. To get these combinations in a minimal number of operations, we co-transfected Cas9 (GFP_2A_Cas9) with sgRNAs against B4Gal-T1, -T2, and -T3 in the first transfection (Table S3, Supporting Information). After single-cell cloning, we carried out deep sequencing to identify clones with exclusively out-of-frame indels in the targeted sequences. In a second round of transfections, we aimed to generate clones with indels in additional B4Gal-T target genes. Therefore we co-transfected GFP_2A_Cas9 with sgRNAs against B4Gal-T1 and -T4 into a clone with confirmed indels in B4Gal-T2 and -T3 (**Table 1**). In our study, a total of 109 potential deletion clones were deep sequenced for genomic indels in the targeted regions (Table S5, Supporting Information). Thereof, clones with in-frame indel or indel frequency <98% were discarded. We expanded clear single- and multi-KO clones of 1–4 targets. Next, we isolated multiple independent clones for each genotype to study true

Table 1. Overview of sgRNA/Cas9 transfections and generated cell lines.

			Target with indel size [bp]			
Parental cell line	Clone name	Transfected with sgRNA against target	B4Gal-T1	B4Gal-T2	B4Gal-T3	B4Gal-T4
CHO-S WT	WT ctr	B4Gal-T1,-T2,-T3				
	T2-3-KO	B4Gal-T1,-T2,-T3		+1	−5	
	T1-3-KO	B4Gal-T1,-T3	+1		+1	
	T3-KO A	B4Gal-T1,-T2,-T3			+1	
	T3-KO B	B4Gal-T1,-T3			+1/−10	
	T1-2-KO A	B4Gal-T1, -T2	+1	+1		
	T1-2-KO B	B4Gal-T1, -T2	+1	−19		
	T1-2-KO C	B4Gal-T1, -T2	+1	+1		
	T1-2-3-KO A	B4Gal-T1,-T2,-T3	+1	−2/−1	+1	
T2-3 KO	T1-2-3-KO B	B4Gal-T1	+1	+1	−5	
	T1-2-3-KO C	B4Gal-T1	+1	+1	−5	
	T1-2-3-KO D	B4GalT-T1	+1	+1	−5	
	T1-2-3-4-KO B	B4Gal-T1, -T4	+1	+1	−5	+1
	T1-2-3-4-KO E	B4Gal-T1, -T4	+2	+1	−5	−1
	T1-2-3-4-KO H	B4Gal-T1, -T4	−1	+1	−5	−1
	T2-3-4-KO H	B4Gal-T1, -T4		+1	−5	−13
	T2-3-4-KO K	B4Gal-T1, -T4		+1	−5	−1
	T2-3-4-KO C	B4Gal-T1, -T4		+1	−5	−13
	T2-3-KO ctr	B4Gal-T1		+1	−5	

The first round of transfections was performed with a CHO-S WT. The T2-3-KO clone was used as a parental cell line for the second transfection round. Values are generated indels in bp for each target confirmed by deep sequencing.

biological replicates of the phenotypes, in total 17 clones (Table 1). Two control clones (WT ctr and T2-3-KO), which showed no additional insertion or deletion after each round of transfection, were additionally selected to investigate the impact of transfection and subcloning on growth and *N*-glycan profile.

3.2. Effect on Growth From Different B4Gal-T-KOs and Indel Combinations

The overall aim of our study is to provide a CHO-S platform to produce recombinant proteins with agalactosylated *N*-glycans. Engineering cells toward G0-glycans leads to altered *N*-glycans on the recombinant protein and also on total host cell glycoproteins. As cell growth performance is a substantial factor for industrial protein production platforms, we first evaluated whether decreased *N*-glycan galactosylation influences CHO cell growth. During shake flask batch experiment, the WT ctr clone was identified to have similar growth and viability compared to CHO-S WT (**Figure 1**). Double-KO of B4Gal-T1 and -T3 (T1-3-KO) indicated decreased growth (Figure 1A), whereas the two clones with frame-shifts in B4Gal-T3 were not influenced in cell growth (Figure 1A). Compared to CHO-S WT and WT ctr, T2-3-KO clone revealed similar growth and T2-3-KO ctr clone exhibited decreased growth (Figure 1B). The four triple-KO clones with frame-shifts in B4Gal-T1, -T2, and -T3 (T1-2-3-KO) and the three T1-2-3-4-KO mutants had decreased growth compared to CHO-S WT (Figure 1B and C). For the three T2-3-4-KO and three T1-2-KO clones, we observed a high diversity in growth between the clones (Figure 1D and E). We additionally calculated the IVC for the different clones over the 7 days of the batch experiment. Compared to CHO-S WT and WT ctr, averages of clone groups without combinatorial disruption of B4Gal-T1 and -T3 had no change in IVC (Figure S4, Supporting Information). In contrast, the IVC from clones harboring a combination of disrupted B4Gal-T1 and -T3 (T1-2-3-KO, T1-3-KO, T1-2-3-4-KO) was slightly decreased.

3.3. Effects of B4Gal-T-KOs on Secretome *N*-Glycan Profiles

To examine the contribution of the targeted B4Gal-Ts on galactosylation of the different *N*-glycan branches, we analyzed the remaining levels of *N*-glycan galactosylation on total secreted proteins in 17 clones with combinatorial disruption of B4Gal-Ts. As presented in Figure S1, Supporting Information, the complex bi-antennary di-sialylated *N*-glycan structure (A2FG2S2) was the major structure within the CHO-S WT secretome. Notably, in CHO-S WT cells, only one minor peak (0.7%) of G0-*N*-glycans could be annotated (Figure S1, Supporting Information). T3-KO A, T2-3-KO, and T2-3-4-KO clones showed a *N*-glycan pattern with minor differences compared to CHO-S WT, and G0 structures were only present in T2-3-KO and T2-3-4-KO clones (**Figure 2** and Figure S1, Supporting Information). In contrast,

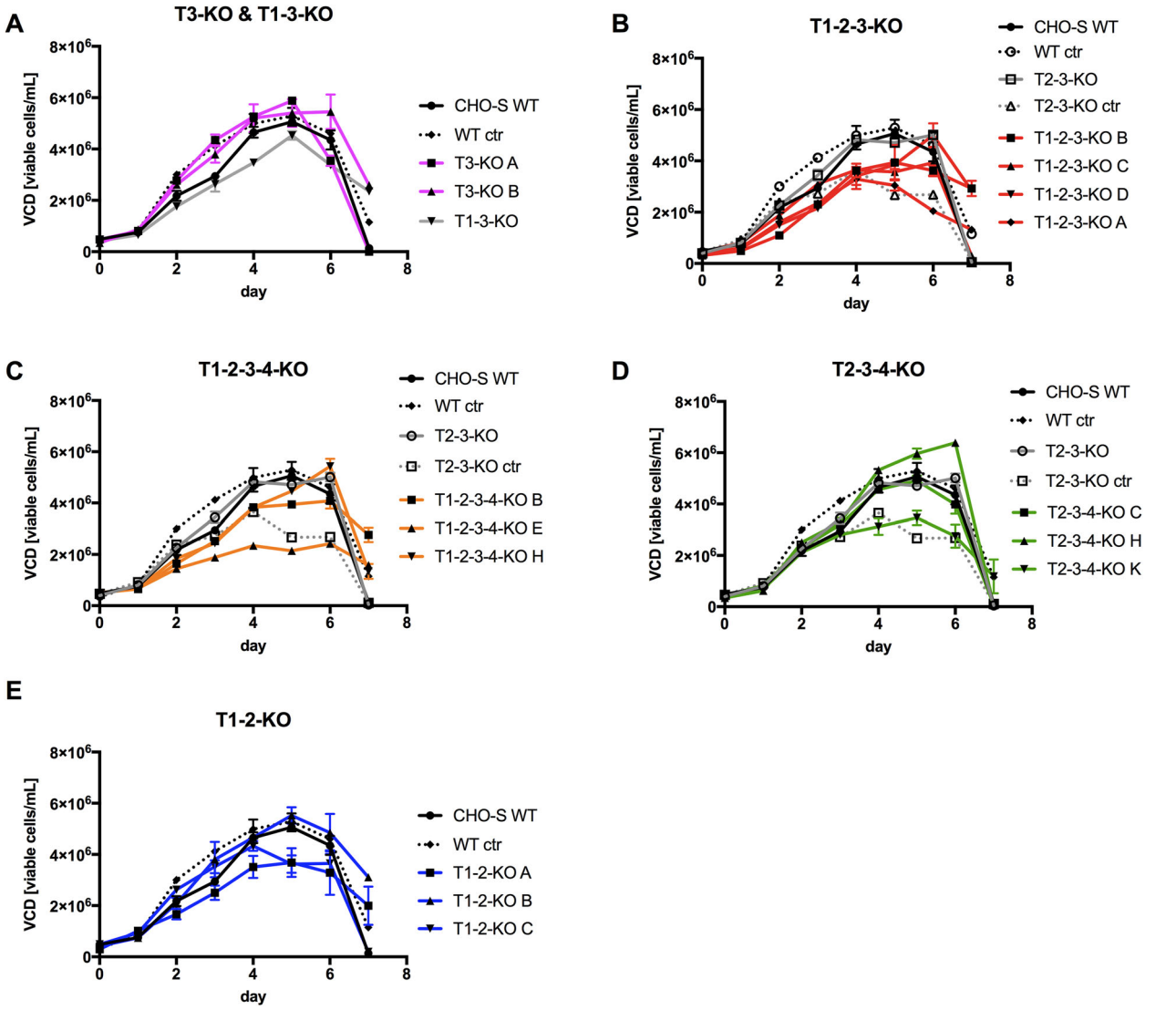


Figure 1. Growth profiles of expanded clones in batch cultivation. A total of 30 mL volume batch cultivation with VCDs for the duration of 7 days after sampling every 24 h ($n = 2$). Seeding was performed at 3.0×10^5 cells mL^{-1}, and error bars indicate range of shake flask duplicates. A) T3-KO & T1-3-KO; B) T1-2-3-KOs; C) T1-2-3-4-KOs; D) T2-3-KOs; E) T1-2-KOs.

the T1-3-KO clone exhibited a total of ≈65% G0 structures and ≈10% G1 *N*-glycans (Figure 2). Indels in B4Gal-T1 and -T2 resulted in the absence of G4 forms, reduced G3 and G2 forms, and increased G1 and G0 proportions (Figure 2). Additionally, we could only annotate ≈1% galactosylated *N*-glycan structures in the secretomes of T1-2-3-KO and T1-2-3-4-KO clones (Figure 2). The major *N*-glycan structures of T1-2-3-KO and T1-2-3-4-KO clones were A2FG0, A3FG0 and A4FG0 (Figure S1, Supporting Information). Furthermore, the additional B4Gal-T4 indel in T1-2-3-4-KO clones did not increase G0 proportions or eliminate G1 *N*-glycans when compared to T1-2-3-KO cell lines (Figure 2). Altogether, disruption of B4Gal-T2 in conjunction with B4Gal-T1 and -T3 decreased the galactosylated secretome *N*-glycan proportion from ≈10% (T1-3-KO) down to ≈1% (T1-2-3-KOs) with A2FG0 as the dominating *N*-glycan structure (Figure S1, Supporting Information). The role of B4Gal-T2 in *N*-glycan galactosylation has previously not been studied in exact terms. To address this, we compared sets of two clones differing in their genotype by the KO of B4Gal-T2 (Figure 2). The occurrence of agalactosylation without (clone T3-KO A and T1-3-KO) and with additional KO of B4Gal-T2 (clones T1-2-3-KO, T2-3-KO, and T2-3-KO ctr) was analyzed. We conclude that the additional disruption of B4Gal-T2 stacked on T3-KO or T1-3-KO increased the proportion of G0-*N*-glycans by ≈3–10% (Figure 2).

3.4. Tailored Rituximab and EPO *N*-Glycosylation After B4Gal-T-Double and Triple-KOs

To investigate if engineered secretome *N*-glycans will also be represented on selected therapeutic proteins, we transiently expressed rituximab and EPO in CHO-S WT and KO clones T3-KO A, T2-3-KO, T1-3-KO, and T1-2-3-KO A and analyzed the resulting *N*-glycan structures upon purification. CHO-S WT,

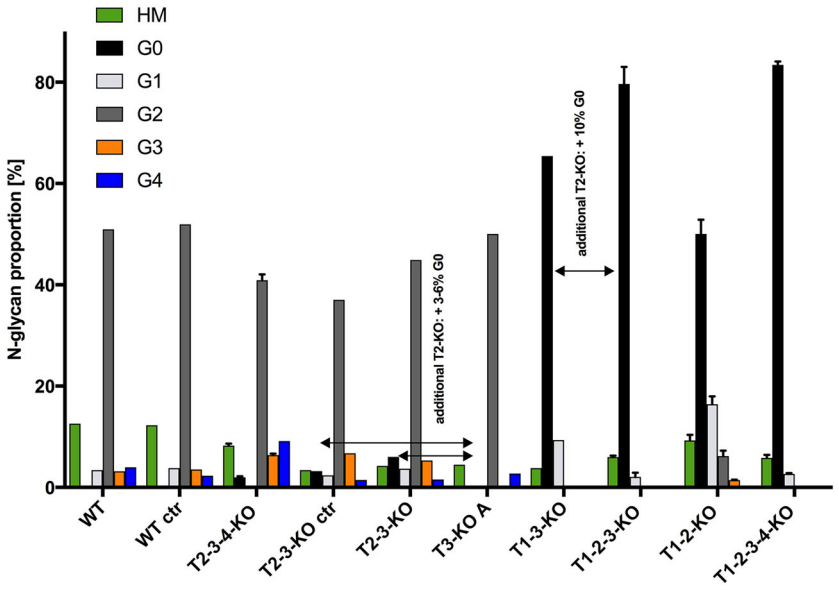


Figure 2. Secretome *N*-glycan profile of generated B4Gal-T-KO clones. *N*-glycan secretome analysis from batch cultivation of parental cell lines and KO cell lines harvested after 5 days of cultivation and normalized to area under the curve (AUC) of total agalactosylated (G0), mono-galactosylated (G1), bi-galactosylated (G2), tri-galactosylated (G3), tetra-galactosylated (G4), and high-mannose (HM) *N*-glycan peaks per cell line. Increase of G0-proportion is given in % after additional B4Gal-T2-KO in T2-3-KO and T1-2-3-KO compared to T3-KO A and T1-3-KO, respectively. Where present, error bars indicate SD of three clones (T1-2-3-4 KO, T2-3-4-KO, and T1-2-KO) or four clones (T1-2-3-KO).

clone T3-KO A, and T2-3-KO displayed comparable rituximab *N*-glycan profiles with G0 and G1 as prevalent structures with both ≈40% of total rituximab *N*-glycans (**Figure 3**A and Figure S2, Supporting Information). In contrast, rituximab purified from clones T1-2-3-KO A and T1-3-KO clones was mostly *N*-glycosylated by bi-antennary G0 structures. As presented in Figure 3B, double-KO of B4Gal-T1 and -T3 (T1-3-KO) resulted in higher G0-*N*-glycan proportions on rituximab (≈84%) than in clone T1-2-3-KO A (≈68%). In both clones, the bi-antennary, G0-galactosylated A2FG0 was clearly the main structure. However, HM, A2G0 and A2FG1 *N*-glycans were also annotated, where T1-2-3-KO A revealed higher HM proportions than other cell lines. Cell growth after rituximab transfection was comparable between CHO-S WT, WT ctr, T3-KO A, and T1-3-KO (Figure S2, Supporting Information), whereas clones T2-3-KO and T1-2-3-KO A revealed increased viable cell concentrations on day 3. For transiently expressed EPO, the *N*-glycan profiles of CHO-S WT, T3-KO A, and T2-3-KO are similar where annotated *N*-glycan structures predominantly harbor ≥4 galactose residues; however, G0 forms are not present in EPO from CHO-S WT (Figure 3C and Figure S3, Supporting Information). In contrast, double-KO of B4Gal-T1 and -T3 resulted in increased G0 proportions (≈72%), whereas G3- and G4-glycans could not be identified. Analyzing *N*-glycan structures of EPO from the triple-KO clone T1-2-3-KO A, we could only annotate agalactosylated and mono-galactosylated *N*-glycans (Figure S3, Supporting Information). Overall, disruption of B4Gal-T1 and -T3 with or without additional disruption of B4Gal-T2 resulted in rituximab with ≈2–3% galactosylated *N*-glycans. Single disruption of B4Gal-T3 or disruption of both B4Gal-T2 and -T3, did not change rituximab *N*-glycosylation. However, disruption of B4Gal-T2 in addition to indels in B4Gal-T1 and -T3 increased the G0 *N*-glycan proportion of transiently expressed EPO from ≈72% to ≈91%.

4. Discussion

Targeting multiple genes in one transfection with CRISPR/Cas9 is a time-saving method to generate clones with different indel combinations in several genes.[13] However, clones often have in-frame indels, which may not disrupt the gene(s).[36] First, we co-transfected with sgRNAs against a combination of B4Gal-T1, -T2, and -T3. In a second round of transfection, we built up triple-KO (T1-2-3-KO and T2-3-4-KO) and quadruple KO clones (T1-2-3-4-KO) based on transfections of the T2-3-KO cell line. Although it is faster, a limitation of this multiplexing method is that not all desired KO combinations might appear after deep sequencing of single-cell clones. An alternative approach would be to use two sgRNAs per target to remove major parts of target DNA sequences from the genome. The double-cut approach is in general less efficient than single cut and furthermore complicates multiplexing. However, in-frame indels become less of a concern.[37]

The effects of B4Gal-T disruptions on cell growth and the glycosylation of total secreted proteins have to our knowledge not been studied in details previously. In a previous study, cell growth was investigated for one clone with a combinatorial disruption of B4Gal-T1 and FUT8.[33] Here, we aimed to assess the impact of B4Gal-T indels on cell growth and *N*-glycosylation in groups of clones with the same combination of indels to additionally address clonal variation. After disrupting the four targets, we observed that clones with indels in B4Gal-T1 and -T3 have decreased IVC when compared to CHO-S WT and WT ctr (Figure 1 and Figure S4, Supporting Information). The reduced IVC in T1-3-KO, T1-2-3-KO, and T1-2-3-4-KO could be associated to the high G0-*N*-glycan proportions of their secretome (Figure 2) or be linked to clonal variation, which is known to be challenging when working with CHO cells.[38] However, glycosylation plays a main role in cell–cell communication via, for example, endocytosis, receptor activation, and cell adhesion,[39] and glycosylation engineering, therefore, might impact cultivation performance. We also report heterogeneous cell growth of clones within the generated indel combination groups. This might also be a result of clonal variation after subcloning or due to off-target effects. While subcloning did not influence growth of the WT ctr clone, subcloning of T2-3-KO lead to decreased growth of the T2-3-KO ctr (Figure 1).

We investigated if disruption of B4Gal-T1, -T2, and -T3 in CHO-S cells is sufficient to produce predominantly agalactosylated proteins and if additional disruption of B4Gal-T4 is of any

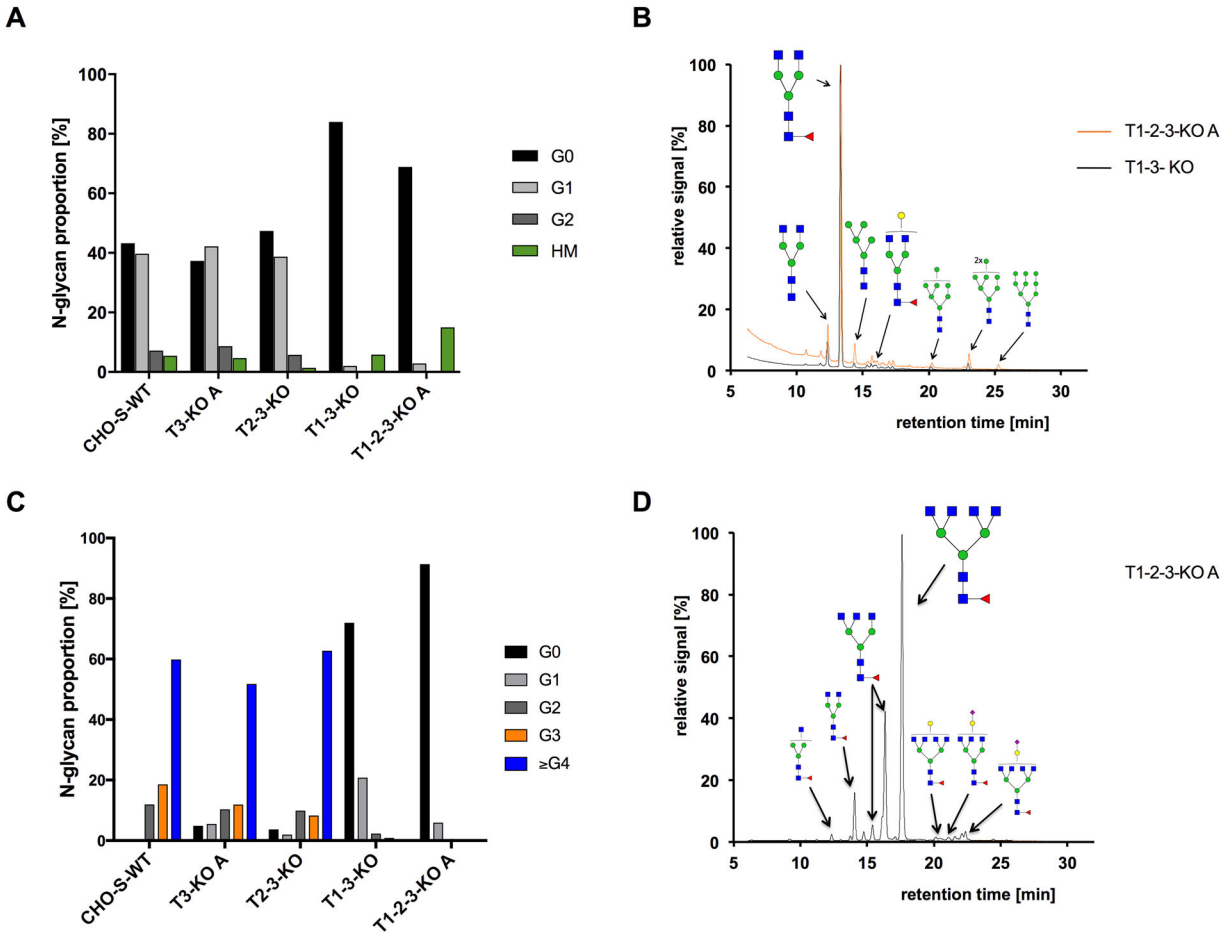


Figure 3. Rituximab and EPO *N*-glycosylation profiles in WT and B4Gal-T KO cell lines after transient transfection. A) Comparison of rituximab *N*-glycans purified out of pooled supernatants within shake flask duplicates from CHO-S WT, T3-KO A, T2-3-KO, T1-2-3-KO A, and T1-3-KO with *N*-glycan proportions of agalactosylated (G0), mono-galactosylated (G1), bi-galactosylated (G2), and high-mannose structures (HM) normalized to AUC of total *N*-glycan peaks per clone. B) Detailed *N*-glycan profiles of rituximab purified out of pooled supernatants within shake flask duplicates from T1-2-3-KO A (orange line) and T1-3-KO (black line) after HPLC histogram annotation via MS. C) Comparison of EPO *N*-glycans purified out of pooled supernatants within shake flask duplicates from CHO-S WT, T3-KO A, T2-3-KO, T1-2-3-KO A, and T1-3-KO with *N*-glycan proportions of agalactosylated (G0), mono- (G1), bi- (G2), tri- (G3) and greater or equal tetra-galactosylated structures (≥G4) normalized to AUC of total *N*-glycan peaks per clone. D) Detailed *N*-glycan profile of EPO purified out of pooled supernatants within shake flask duplicates from T1-2-3-KO A.

benefit for decreased galactosylation. In contrast to a previous study, which suggested B4Gal-T1–4 to all be active in *N*-glycan galactosylation,[30] our results indicate that B4Gal-T1, -T2, and -T3 are the most active B4Gal-Ts in the *N*-glycosylation pathway of CHO-S cells and that B4Gal-T4 has very little or no contribution to galactosylation of *N*-glycans in CHO-S cells. The lack of *N*-glycosylation activity of B4Gal-T4 in our work supports another study where B4Gal-T4 was reported to be active in the galactosylation of mucin-type core 2 branching in the O-glycosylation pathway.[31] Furthermore, B4Gal-T5, -T6, and -T7 (and potentially unknown B4Gal-Transferases) in sum contribute only up to ≈3% *N*-glycan galactosylation of the secretome, as seen in Figure 2. Our results indicate that subcloning had no impact on secretome *N*-glycosylation as the WT ctr and T2-3-KO ctr clones showed comparable *N*-glycan structures to their parental cell lines in the batch cultivation (Figure 2).

As the T2-3-KO clone still produced G1, G2, G3, and G4 structures and all KO cell lines with indels in B4Gal-T1 lack G4 *N*-glycans, B4Gal-T1 is very likely capable of transferring galactose to all four branches. Therefore, we suggest that B4Gal-T1 is the most active *N*-glycan processing B4Gal-T within the family of β-1,4-galactosyltransferases of CHO-S cells, which is in line with previous work in another CHO cell line.[33] We expected to reveal branch specificities for all four targeted B4Gal-Ts. Due to low galactosylation activities of B4Gal-T2, -T3, and -T4, we can only conclude that B4Gal-T1 can galactosylate all four antennas (Figure 2) and that its branch preference needs to be explored further.

Moreover, we analyzed that B4Gal-T2 activity contributes to ≈3–10% of *N*-glycan galactosylation (Figure 2). Since single B4Gal-T3-KO did not decrease galactosylation (Figure 2), we suggest that B4Gal-T3 has only a minor role in CHO-S *N*-glycosylation or that its disrupted *N*-glycan transferase

function can be compensated by B4Gal-T1 and -T2 activity in the T3-KO clone.

The remaining level of rituximab galactosylation of the CHO-S-derived clone T1-3-KO (≈2–3%) is comparable, yet slightly higher to another study where decreased rituximab galactosylation (≈1%) was achieved by knocking out B4Gal-T1 and -T3 in CHO-K1-derived cell lines.[33] This difference in remaining *N*-glycan galactosylation could be due to differences in the *N*-glycan pathways of the cell lines used (CHO-S versus CHO-K1)[33] or due to clonal variation. Within our triple-KO cell line T1-2-3-KO A, we also noticed a significant amount of high-mannose (HM) structures on transiently expressed rituximab (Figure 3A and B). HM structures are a critical quality attribute within biopharmaceutical protein production[40] and can accumulate during cell culture performance. Process design and genetic engineering could be two possibilities to overcome accumulated HM structures, which might represent proteins accumulated in the Golgi-situated *N*-glycan machinery after disrupting Golgi-residing B4Gal-T1, -T2, and -T3. This disruption might cause increased traffic and residence time of secretome proteins in the Golgi lumen without being further processed by glycosyltransferases. Recent studies displayed increased processing of *N*-glycans after overexpression of Mgat4 and Mgat5, which could result in lower HM proportions on glycoproteins.[41]

For glycoproteins harboring tri- or tetra-antennary *N*-glycans, as is the case for EPO, KO of B4Gal-T1 and -T3 is not sufficient to produce mainly agalactosylated glycoproteins (Figure 3B with ≈20% EPO galactosylation in T1-3-KO). In comparison, rituximab expressed in clone T1-3-KO resulted in only ≈3% galactosylated structures (Figure 3A). Therefore, we propose that bi-antennary *N*-glycosylated proteins as rituximab can be produced with mostly agalactosylated *N*-glycans after double-KO of B4Gal-T1 and -T3. However, tri- and tetra-antennary *N*-glycosylated secretome proteins as EPO additionally need KO of B4Gal-T2 to be predominantly agalactosylated. For transiently expressed EPO in CHO-K1-derived cells with triple-KO of B4Gal-T1, -T2, and -T3 the proportions of galactosylated *N*-glycans were found to be ≈4% in an earlier study.[33] In our study, we annotated ≈6% galactosylated *N*-glycans on transiently expressed EPO from the CHO-S-derived triple-KO T1-2-3-KO A (Figure 3C). Although these results indicate similar effects on galactosylation of EPO after disruption of two identical gene targets, deviations could be related to differences between CHO-K1 and CHO-S expression levels of nontargeted B4Gal-T isoforms.

Therefore, we suggest that engineering cells with non-galactosylated *N*-glycans on a secretome level in CHO-S WT is a promising strategy toward producing G0-IgG1 and G0-EPO at a later stage. Despite the divergent gene expression levels between different CHO cell lines[42] this engineering strategy is suitable for both CHO-K1[33] and CHO-S-derived cell lines as utilized in our work. In the presented study the triple-KO with ≈1% galactosylated structures on the secretome also showed predominantly agalactosylated *N*-glycans on transiently expressed rituximab with only ≈3% galactosylated *N*-glycans and on transiently expressed EPO with remaining ≈6% galactosylated *N*-glycan structures.

In summary, our study presents the necessity of disrupting the three genes, B4Gal-T1, -T2, and -T3, to produce predominantly agalactosylated secretome proteins, rituximab and EPO in CHO-S cells. The possibility to engineer tri- and tetra-antennary G0 *N*-glycans, which are naturally not produced in CHO-S WT cells (Figure S1, Supporting Information), is presented. Prior engineering of secretome *N*-glycans in a WT cell gives rise to the flexibility of expressing several different model proteins in the engineered cell line at a later stage. Such model proteins might include already marketed antibodies or other therapeutic proteins. With our cell platform that prevalently glycosylates proteins with G0-*N*-glycans, we demonstrate an alternative to galactosidase treatment of recombinant proteins to investigate further beneficial in vitro and in vivo characteristics based on tailored G0 *N*-glycosylation profiles.

Abbreviations

Asn, Asparagine; AUC, area under curve; B4Gal-T, β-1,4-galactosyltransferase; Cas9, CRISPR-associated protein 9; CHO, Chinese hamster ovary; CRISPR, clustered regularly interspaced short palindromic repeats; EPO, erythropoietin; FACS, fluorescence-activated cell sorting; Fc, fragment crystallizable; FcγRIIIa, Fc-gamma receptor III a; FUT8, alpha-(1,6)-fucosyltransferase; G0, agalactosylated; GlcNAc, *N*-Acetylglucosamine; HM, high-mannose; HPC4, human protein C4; IgG, immunoglobulin G; indel, insertion or deletion; mAb, monoclonal antibody; sgRNA, single guide RNA; TSTA3, tissue-specific transplantation antigen P35B; UDP-Gal, uridine diphosphate galactose; VCD, viable cell density; WT, wild type.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

T.A. and A.H.H. contributed equally to this publication. The authors thank Sara Petersen Bjørn, Bjørn Voldborg, Johnny Arnsdorf, Yuzhou Fan, and Patrice Menard for valuable guidance and support. The authors thank Karen Katrine Brøndum, Nachon Charanyanonda Petersen, Karoline Schousboe Fremming, and Zulfiya Sukhova for excellent technical assistance with the FACS and MiSeq library preparation; Helle Munck Petersen for assistance with the protein purification; Anna Koza and Mads Valdemar Anderson for assistance with the MiSeq analysis; and Johan Rockberg (KTH, Sweden) for providing the HPC-4 sequence. The Novo Nordisk Foundation (NNF10CC1016517) supported this work. T.A., H.F.K., and M.R.A. are receiving funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No. 642663. A.H.H., M.R.A., H.F.K., and T.A. planned the experiments. T.A. performed the experimental work and wrote the manuscript. A.H.H. performed the *N*-glycan analysis, and S. K. conducted the protein purifications. A.H.H., M.R.A., H.F.K., S.K., and G.M.L. guided the project, contributed to experimental design, and commented and corrected the manuscript.

Conflict of Interest

The authors declare no commercial or financial conflict of interest.

Keywords

Chinese hamster ovary cells, CRISPR/Cas9, erythropoietin, glycoengineering, multiplexing, *N*-glycosylation, rituximab

Received: February 17, 2018
Revised: May 4, 2018
Published online:

[1] J. Zhu, *Biotechnol. Adv.* **2012**, *30*, 1158.
[2] J. Knaeblein, *Modern Biopharmaceuticals: Recent Success Stories*. (Ed. J. Knaeblein), Wiley-VCH, Weinheim **2013** Ch. 1.
[3] P. H. Lai, R. Everett, F. F. Wang, T. Arakawa, E. Goldwasser, *J. Biol. Chem*. **1986**, *261*, 3116.
[4] M. Dalziel, M. Crispin, C. N. Scanlan, N. Zitzmann, R. A. Dwek, *Science* **2014**, *343*, 1235681.
[5] M. F. Jennewein, G. Alter, *Trends Immunol*. **2017**, *38*, 358.
[6] Y. Kanda, T. Yamada, K. Mori, A. Okazaki, M. Inoue, K. Kitajima-Miyama, R. Kuni-Kamochi, R. Nakano, K. Yano, S. Kakita, K. Shitara, M. Satoh, *Glycobiology* **2007**, *17*, 104.
[7] G. Walsh, *Nat. Biotechnol*. **2014**, *32*, 992.
[8] O. Montacir, H. Montacir, M. Eravci, A. Springer, S. Hinderlich, A. Saadati, M. K. Parr, *J. Pharm. Biomed. Anal*. **2017**, *140*, 239.
[9] S. Iida, H. Misaka, M. Inoue, M. Shibata, R. Nakano, N. Yamane-Ohnuki, M. Wakitani, K. Yano, K. Shitara, M. Satoh, *Clin. Cancer Res*. **2006**, *12*, 2879.
[10] M. Satoh, S. Iida, K. Shitara, *Expert Opin. Biol. Ther*. **2006**, *6*, 1161.
[11] N. Yamane-Ohnuki, M. Satoh, *MAbs* **2009**, *1*, 230.
[12] L. Malphettes, Y. Freyvert, J. Chang, P. Q. Liu, E. Chan, J. C. Miller, Z. Zhou, T. Nguyen, C. Tsai, A. W. Snowden, T. N. Collingwood, P. D. Gregory, G. J. Cost, *Biotechnol. Bioeng*. **2010**, *106*, 774.
[13] L. M. Grav, J. S. Lee, S. Gerling, T. B. Kallehauge, A. H. Hansen, S. Kol, G. M. Lee, L. E. Pedersen, H. F. Kildegaard, *Biotechnol. J*. **2015**, *10*, 1446.
[14] S. Louie, B. Haley, B. Marshall, A. Heidersbach, M. Yim, M. Brozynski, D. Tang, C. Lam, B. Petryniak, D. Shaw, J. Shim, A. Miller, J. B. Lowe, B. Snedecor, S. Misaghi, *Biotechnol. Bioeng*. **2017**, *114*, 632.
[15] F. Nimmerjahn, R. M. Anthony, J. V. Ravetch, *Proc. Natl. Acad. Sci. U. S. A*. **2007**, *104*, 8433.
[16] R. Malhotra, M. R. Wormald, P. M. Rudd, P. B. Fischer, R. A. Dwek, R. B. Sim, *Nat. Med*. **1995**, *1*, 237.
[17] A. Wright, S. L. Morrison, *J. Immunol*. **1998**, *160*, 3393.
[18] T. A. Millward, M. Heitzmann, K. Bill, U. Längle, P. Schumacher, K. Forrer, *Biologicals* **2008**, *36*, 41.
[19] A. M. Goetze, Y. D. Liu, Z. Zhang, B. Shah, E. Lee, P. V. Bondarenko, G. C. Flynn, *Glycobiology* **2011**, *21*, 949.
[20] M. E. Ackerman, M. Crispin, X. Yu, K. Baruah, A. W. Boesch, D. J. Harvey, A. S. Dugast, E. L. Heizen, A. Ercan, I. Choi, H. Streeck, P. A. Nigrovic, C. Bailey-Kellogg, C. Scanlan, G. Alter, *J. Clin. Invest*. **2013**, *123*, 2183.
[21] C. Lood, M. Allhorn, R. Lood, B. Gullstrand, A. I. Olin, L. Rönnblom, L. Truedsson, M. Collin, A. A. Bengtsson, *Arthritis Rheum*. **2012**, *64*, 2698.
[22] G. Dekkers, R. Plomp, C. A. M. Koeleman, R. Visser, H. H. von Horsten, V. Sandig, T. Rispens, M. Wuhrer, G. Vidarsson, *Sci. Rep*. **2016**, *6*, 36964.
[23] X. Xu, H. Nagarajan, N. E. Lewis, S. Pan, Z. Cai, X. Liu, W. Chen, M. Xie, W. Wang, S. Hammond, M. R. Andersen, N. Neff, B. Passarelli, W. Koh, H. C. Fan, J. Wang, Y. Gui, K. H. Lee, M. J. Betenbaugh, S. R. Quake, I. Famili, B. O. Palsson, J. Wang, *Nat. Biotechnol*. **2011**, *29*, 735.
[24] M. Amado, R. Almeida, T. Schwientek, H. Clausen, *Biochim. Biophys. Acta Gen. Subj*. **1999**, *1473*, 35.
[25] K. Furukawa, T. Sato, *Biochim. Biophys. Acta* **1999**, *1473*, 54.
[26] I. Van Die, A. Van Tetering, W. E. C. M. Schiphorst, T. Sato, K. Furukawa, D. H. Van Den Eijnden, *FEBS Lett*. **1999**, *450*, 52.
[27] J. Lee, S. Sundaram, N. L. Shaper, T. S. Raju, P. Stanley, *J. Biol. Chem*. **2001**, *276*, 13924.
[28] T. Okajima, K. Yoshida, T. Kondo, K. Furukawa, *J. Biol. Chem*. **1999**, *274*, 22915.
[29] R. Almeida, *J. Biol. Chem*. **1999**, *274*, 26165.
[30] S. Guo, T. Sato, K. Shirane, K. Furukawa, *Glycobiology* **2001**, *11*, 813.
[31] M. Ujita, J. McAuliffe, T. Schwientek, R. Almeida, O. Hindsgaul, H. Clausen, M. Fukuda, *J. Biol. Chem*. **1998**, *273*, 34843.
[32] M. Asano, K. Furukawa, M. Kido, S. Matsumoto, Y. Umesaki, N. Kochibe, Y. Iwakura, *EMBO J*. **1997**, *16*, 1850.
[33] Z. Yang, S. Wang, A. Halim, M. A. Schulz, M. Frodin, S. H. Rahman, M. B. Vester-Christensen, C. Behrens, C. Kristensen, S. Y. Vakhrushev, E. P. Bennett, H. H. Wandall, H. Clausen, *Nat. Biotechnol*. **2015**, *33*, 842.
[34] P. G. Slade, M. Hajivandi, C. M. Bartel, S. F. Gorfien, *J. Proteome Res*. **2012**, *11*, 6175.
[35] C. Ronda, L. E. Pedersen, H. G. Hansen, T. B. Kallehauge, M. J. Betenbaugh, A. T. Nielsen, H. F. Kildegaard, *Biotechnol. Bioeng*. **2014**, *111*, 1604.
[36] A. Nicolas, C. Lucchetti-Miganeh, R. Yaou, J. Kaplan, J. Chelly, F. Leturcq, F. Barloy-Hubler, E. Le Rumeur, *Orphanet. J. Rare Dis*. **2012**, *7*, 45.
[37] V. Schmieder, N. Bydlinski, R. Strasser, M. Baumann, H. F. Kildegaard, V. Jadhav, N. Borth, *Biotechnol. J*. **2018**, *13*, 1700211.
[38] B. N. Vcelar, V. Jadhav, M. Melcher, N. Auer, A. Hrdina, R. Sagmeister, K. Heffner, A. Puklowski, M. Betenbaugh, T. Wenger, F. Leisch, M. Baumann, *Biotechnol. Bioeng*. **2017**, *115*, 165.
[39] K. Ohtsubo, J. D. Marth, *Cell* **2006**, *126*, 855.
[40] M. Yu, D. Brown, C. Reed, S. Chung, J. Lutman, E. Stefanich, A. Wong, J. P. Stephan, R. Bayer, *MAbs* **2012**, *4*, 475.
[41] B. Yin, Y. Gao, C. yu Chung, S. Yang, E. Blake, M. C. Stuczynski, J. Tang, H. F. Kildegaard, M. R. Andersen, H. Zhang, M. J. Betenbaugh, *Biotechnol. Bioeng*. **2015**, *112*, 2343.
[42] N. E. Lewis, X. Liu, Y. Li, H. Nagarajan, G. Yerganian, E. O'Brien, A. Bordbar, A. M. Roth, J. Rosenbloom, C. Bian, M. Xie, W. Chen, N. Li, D. Baycin-Hizal, H. Latif, J. Forster, M. J. Betenbaugh, I. Famili, X. Xu, J. Wang, B. O. Palsson, *Nat. Biotechnol*. **2013**, *31*, 759.

CHAPTER 4

Glyco-engineered CHO cell lines producing alpha-1-antitrypsin and C1 esterase inhibitor with fully humanized N-glycosylation profiles

As described in the previous chapters, CRISPR/Cas9 multiplexing in CHO can be used for the design of certain N-glycan profiles on therapeutic proteins. In Chapter 4 we (i) analyzed the N-glycan structures of the two marketed drugs Prolastin® (A1AT) and Cinryze® (C1INH), which are predominantly sold as purifications from human plasma, to then (ii) generate a cell line with disruptions in ten gene targets by CRISPR/Cas9 multiplexing for (iii) the expression of a human glycosyltransferase which enables the production of recombinant A1AT and C1INH with N-glycan profiles and *in vitro* activities similar to the counterparts from human plasma. By providing this cell line, we contribute to the strategy recommended by the Medical and Scientific Advisory Council to replace plasma-derived therapeutics with more safe and scalable recombinant products for the treatment of diseases[1].

1 National Hemopilia Foundation, https://www.hemophilia.org/Researchers-Healthcare-Providers/Medical-and-Scientific-Advisory-Council-MASAC/MASAC-Recommendations/MASAC-Recommendations-Regarding-Standards-of-Service-for-Pharmacy-Providers-of-Clotting-Factor-Concentrates-for-Home-Use-to-Pat. 2014.

Glyco-engineered CHO cell lines producing alpha-1-antitrypsin and C1 esterase inhibitor with fully humanized N-glycosylation profiles

Thomas Amann[1], Anders Holmgaard Hansen[1], Stefan Kol[1], Henning Gram Hansen[1], Johnny Arnsdorf[1], Saranya Nallapareddy[1], Bjørn Voldborg[1], Gyun Min Lee[1,2], Mikael Rørdam Andersen[3], Helene Faustrup Kildegaard[1]

[1]The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark

[2]Department of Biological Sciences, KAIST, Daejeon, Republic of Korea

[3]Department of Biotechnology and Biomedicine, Technical University of Denmark, Kgs. Lyngby, Denmark

---

Abbreviations:

A1AT, alpha-1-antitrypsin; **AATD**, alpha-1-antitrypsin deficiency; **AUC**, area under curve; **B3gnt2**, UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 2; **Cas9**, CRISPR-associated protein 9; **C1INH**, C1 esterase inhibitor; **CHO**, Chinese hamster ovary; **CRISPR**, clustered regularly interspaced short palindromic repeats; **FACS**, fluorescence-activated cell sorting; **FITC**, Fluorescein isothiocyanate; **FUT8**, alpha-(1,6)-fucosyltransferase; **GLUL**, glutamate-ammonia ligase; **HAE**, hereditary angioedema; **HM**, high–mannose; **indel**, insertion or deletion; **IVC**, integral of viable cells; **KO**, knock-out; **mAb**, monoclonal antibody; **MGAT4A**, mannosyl (alpha-1,3-)-glycoprotein beta-1,4-N-acetylglucosaminyltransferase isozyme A; **MGAT4B**, mannosyl (alpha-1,3-)-glycoprotein beta-1,4-N-acetylglucosaminyltransferase isozyme B; **MGAT5**, mannosyl (alpha-1,6-)-glycoprotein beta-1,6-N-acetylglucosaminyltransferase; **MSX**, methionine sulfoximine; **sgRNA**, single guide RNA; **SNA**, sambucus nigra agglutinin; **SPPL3**, signal peptide peptidase like 3; **ST3GAL3**, ST3 beta-galactoside alpha-2,3-sialyltransferase 3; **ST3GAL4**, ST3 beta-galactoside alpha-2,3-sialyltransferase 4; **ST3GAL6**, ST3 beta-galactoside alpha-2,3-sialyltransferase 6; **ST6GAL1**, ST6 beta-galactoside alpha-2,6-sialyltransferase 1; **VCD**, viable cell density; **WT**, wild type

Abstract

Recombinant Chinese hamster ovary (CHO) cells are able to provide biopharmaceuticals that are essentially free of human viruses and have N-glycosylation profiles similar, but not identical, to humans. Due to differences in N-glycan moieties, two members of the serpin superfamily, alpha-1-antitrypsin (A1AT) and plasma protease C1 inhibitor (C1INH), are currently derived from human plasma for treating A1AT and C1INH deficiency. Deriving therapeutic proteins from human plasma is generally a cost-intensive process and also harbors a risk of transmitting infectious particles. Recombinantly produced A1AT and C1INH (*rh*A1AT, *rh*C1INH) decorated with humanized N-glycans are therefore of clinical and commercial interest.

Here, we present engineered CHO cell lines producing *rh*A1AT or *rh*C1INH with fully humanized N-glycosylation profiles. This was achieved by combining CRISPR/Cas9-mediated disruption of 10 gene targets with overexpression of human ST6GAL1. We were able to show that the N-linked glyco-structures of *rh*A1AT and *rh*C1INH are homogeneous and similar to the structures obtained from plasma-derived A1AT and C1INH, marketed as Prolastin®-C and Cinryze®, respectively. *rh*A1AT and *rh*C1INH produced in our glyco-engineered cell line showed no detectable differences to their plasma-purified counterparts on SDS-PAGE and had similar enzymatic *in vitro* activity. The work presented here shows the potential of expanding the glyco-engineering toolbox for CHO cells to produce a wider variety of glycoproteins with fully humanized N-glycan profiles. We envision replacing plasma-derived A1AT and C1INH with recombinant versions and thereby decreasing our dependence on human donor blood, a limited and possibly unsafe protein source for patients.

1 Introduction

Chinese hamster ovary (CHO) cells serve an important role in the biotechnology industry as the primary workhorse for the production of recombinant protein therapeutics [1]. Many of these therapeutics are glycoproteins that contain one or more N-glycan and/or O-glycan chains. As N-glycans can potentially affect protein folding, immune regulation, cellular homeostasis and the biological half-life of proteins [2, 3], it is considered a critical quality attribute and much effort has been put forth to improve features of protein N-glycosylation. The production of diverse N-glycan structures is a major contributor to the heterogeneity of protein products derived from CHO cells. The inherent heterogeneity of CHO N-glycan profiles is especially a drawback when one distinct N-glycan structure is desired on the protein product. Two examples of human plasma proteins with distinct, homogeneous N-glycan structures are found within the serpin superfamily, alpha-1-antitrypsin (A1AT)

and C1 esterase inhibitor (C1INH) [4]. Patients with the genetic disorders alpha-1-antitrypsin deficiency (AATD) or hereditary angioedema (HAE-C1INH) have decreased plasma levels of functional A1AT or C1INH, respectively, and are currently treated with prophylactic augmentation therapy of plasma purified A1AT or C1INH[5, 6]. Augmentation therapy is cost intensive [5], and C1INH purified from pooled donor plasma has been associated with hepatitis C virus infections prior to the introduction of virucidal methods[7]. Despite current dedicated virus inactivation steps, cases of Hepatitis G transmission have been reported [8] and non-enveloped viruses can still be transmitted via plasma-derived products [9]. Nevertheless, approved C1INH formulas are concentrates purified from human donors (Berinert®, Cinryze) despite containing undesired protein impurities identified as α1-antichymotrypsin, ceruloplasmin and Factor C3 [6].

Both native human plasma A1AT (*pl*A1AT) and C1INH (*pl*C1INH) possess a N-glycan profile with ~60–80% diantennary, disialylated, non-fucosylated (A2G2S2) structures with human-like alpha-2,6-linked sialic acids. *pl*A1AT has three N-glycosylation sites [4], where natural A2G2S2 structures are not essential for biological activity but enhance *in vivo* half-life and *in vitro* protein stability [10, 11]. *pl*C1INH is thought to be one of the most heavily glycosylated plasma proteins and harbors ten O-linked and six N-linked glycan structures [12]. The six N-linked glycan moieties with A2G2S2 as predominant structure have been shown to be increase serum half-life and are reported to increase *in vivo* efficacy [13–15].

The ability to generate high A2G2S2 N-glycan proportions on *rh*A1AT/*rh*C1INH may be critical for improving product quality. Although efforts have been reported, generating high proportions of A2G2S2 N-glycans has not to our knowledge been published. For instance, CHO cells lack active St6 beta galactoside alpha-2,6-sialyltransferase 1 (ST6GAL1) to cap N-glycans with alpha-2,6-linked sialic acids [16]. *rh*A1AT and *rh*C1INH were produced in various platforms [17-29]. However, these approaches revealed low productivity or the N-glycosylation was far from the profile of *pl*A1AT or *pl*C1INH and therefore cleared rapidly from the human blood making intravenous administration impractical.

Commercially available *rh*C1INH from transgenic rabbits shows activity similar to *pl*C1INH and has decreased virus transmission risk. However, it differs in N-glycosylation profiles from *pl*C1INH and therefore reveals a risk of allergy, a dissatisfactory pharmacokinetic profile and consequently is unlikely to be of use in prophylaxis [30]. Glycosylation-engineering in primary human cells aimed to mimic O-glycan profiles of *pl*C1INH. However, complete sialylation of *rh*C1INH N-glycan structures was not achieved [31].

By combining the publicly available CHO-K1 genome sequence [32], clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9 (CRISPR/Cas9) for multiplexing gene editing [33] and reported CHO glycosyltransferases [34], we aimed to engineer the heterogeneous CHO-S N-glycan profile towards a predominantly non-fucosylated biantennary A2G2 structure. To this end, we made functional knockouts of the GLUL-gene and nine glycosylation-gene targets (10x KO, Suppl. Table 1). We hypothesized that *rh*A1AT and *rh*C1INH produced in this genetic background with parallel co-expression of St6gal1 would display an N-glycan profile similar to *pl*A1AT and *pl*C1INH with predominant A2G2S2 N-glycan structures. We present CHO clones producing *rh*A1AT/*rh*C1INH similar to *pl*A1AT/*pl*C1INH by N-glycan analysis, protein activity, SDS-PAGE and isoelectric focusing. CHO derived *rh*A1AT and *rh*C1INH with fully humanized N-glycan profiles have the potential to replace the cost-intensive and possibly unsafe plasma-based augmentation therapy of AATD and HAE-C1INH patients without compromising activity and N-glycosylation.

2 Materials and methods

2.1 sgRNA, GFP_2A_Cas9 and A1AT/C1INH_ST6GAL1_GLUL plasmid design

GFP_2A_Cas9 and single guide RNA (sgRNA) plasmids were constructed as previously described [33]. The sgRNA target design for MGAT4A, MGAT4B, MGAT5, ST3GAL3, ST3GAL4, ST3GAL6, B3GNT2, FUT8, SPPL3 and GLUL was performed using “CRISPy” [35]. The target sites for the mentioned genes and the oligos for sgRNA cloning are listed in Suppl. Table S1 and Table S2, respectively.

Plasmids for co-expression of A1AT/C1INH and ST6GAL1 were constructed with uracil-specific excision reagent cloning method as previously described [36, 37] (Suppl. Fig. 1). The DNA sequences of the plasmids are listed in Suppl. Table S5.

2.2 Cell cultivation and transfection for genome editing

CHO-S suspension cells were incubated in a humidified incubator at 120 rpm, 37°C, 5% CO_2, passaged to 2-3 x 10^5 cells/mL every 2-3 days and transfected in 6-well plates (BD Biosciences, San Jose, CA) as described previously [33]. The GFP_2A_Cas9 / sgRNA plasmid ratios for each transfection was 1:1 of which the plasmid load of sgRNA was divided equally by the amount of different sgRNAs used per transfection (Suppl. Table S4). To measure FACS sorting efficiency, pmaxGFP® vector (Lonza, Basel, Switzerland) transfection was performed as well. Cells were

harvested for fluorescence-activated cell sorting (FACS) 48 h post transfection.

2.3 Single cell cloning of genome edited cells using FACS

Before FACS, cells were filtered through a 40 µm cell strainer into a FACS-compatible tube. Single fluorescent-positive (GFP) cells were sorted into 384-well plates (Corning, New York, NY) containing 30 µL CD CHO medium supplemented with 8 mM L-glutamine, 1.5% HEPES buffer and 1% Antibiotic-Antimycotic (Gibco, Waltham, MA) per well as described previously [38]. For cell sorting, fluorescent-positive cell populations were gated based on non-transfected WT CHO-S cells. Two weeks after cell sorting cell colonies were moved to 96-well flat-bottom plates (BD Biosciences) and expanded for deep sequencing analysis and batch cultivation.

2.4 Deep sequencing analysis

Confluent colonies from 96-well flat-bottom replicate plates were harvested for genomic DNA extraction. DNA extraction was performed using QuickExtract DNA extraction solution (Epicentre, Illumina, Madison, WI) according to the manufacturer's instruction. The library preparation was based on Illumina 16S Metagenomic Sequencing Library Preparation and deep sequencing was carried out on a MiSeq Benchtop Sequencer (Illumina, San Diego, CA). The protocol for amplifying the targeted genomic sequences, amplicon purification, adapter-PCR and following quality analysis was based on previously published work [33]. PCR primers are presented in Suppl. Table S3.

2.5 Transfection and expression in polyclonal cell lines by applying MSX-selection

Cells were seeded in 250 mL Corning vent cap shake flasks (Sigma-Aldrich) as duplicates with cell densities ~1 x 10^6 cells/mL in 60 mL CD CHO medium supplemented with 8 mM L-glutamine (Life Technologies) and transfected with 75 µg of A1AT-GLUL-St6gal plasmid or 75 µg of C1INH-GLUL-ST6GAL1 plasmid (Suppl. Fig. 1) using FreeStyle™ MAX reagent together with OptiPRO SFM medium (Life Technologies) according to the manufacturer's recommendations. 1µL/mL anti-clumping agent was added 24 h after transfection. pmaxGFP® vector (Lonza) transfection was performed to measure transfection efficiencies. Two days after transfection, cells were transferred into 60 mL CD CHO medium lacking L-glutamine (Life Technologies) and supplemented with 1µL/mL anti-clumping agent and 0 µM, 10 µM, 30 µM or 50 µM MSX (EMD Millipore, Billerica, MA).

Cell densities and viabilities were determined once per day using the NucleoCounter NC-250 Cell Counter (ChemoMetec). The cells were passaged in fresh selection medium every 2-3 days until viability and doubling time reached stable values. Polyclonal cell lines (pools) were seeded in

duplicates at ~1 x 10^6 cells/mL with corresponding MSX concentrations. Cell densities and viabilities were determined once per day and supernatants of the pools were harvested three days after seeding and pooled within duplicates for purification of *rh*A1AT and *rh*C1INH.

2.6 Single cell cloning of cells from polyclonal cell pools using FACS

Non-stained single cells were sorted from pools as described above. For cell sorting, all viable cells were gated for sorting into 384-well plates with L-glutamine-free medium. Two weeks after cell sorting the clones were moved to 96-well flat-bottom plates (BD Biosciences) and expanded to shake flask format in CD CHO medium supplemented with 1µL/mL anti-clumping agent, 25 µM MSX and lacking L-glutamine.

2.7 Screening cell pools and single cell clones for human-like α-2,6-sialic acid linkage formation with lectin staining

For lectin staining of cells, triplicates of 10,000 cells per sample were diluted in 200 µL of 0.22 µm pore size filtered CD CHO medium (Life Technologies) supplemented with 5 µg/mL Hoechst 33342 (Merck, Darmstadt, Germany) and 1 µg/mL Fluorescein isothiocyanate (FITC) labeled *Sambucus nigra* agglutinin (SNA) lectin (Biomol, Hamburg, Germany). After 60 min incubation in the dark at 37°C and 5% CO_2, the cells were washed with 200 µL CD CHO medium and then washed twice with 200 µL phosphate buffered saline (PBS) (300g, 5 min, RT). The samples were resuspended in 200 µl PBS and transferred to 96-well plate for final centrifugation at 300 g for one minute. Percentage of FITC SNA positive cells was determined in a 96-well optical-bottom microplate (Greiner Bio-One, Frickenhausen, Germany) using a Celigo Imaging Cell Cytometer (Nexcelom Bioscience, Lawrence, MA). Cells were identified using the blue channel (Hoechst-positive cells), and the green channel (FITC SNA-positive cells) was used to detect cells with alpha-2,6-sialic acid linkage. A Hoechst/FITC SNA-stained CHO-S WT sample was gated to distinguish between FITC-positive and FITC-negative cells.

2.8 Batch cultivation: cell growth analysis and N-glycosylation profiling

For batch cultivation and N-glycan analysis, cells were seeded at 0.4 x 10^6 cells/mL in 250 mL Corning vent cap shake flasks (Sigma-Aldrich, St. Louis, MI) as duplicates in 60 mL CD CHO medium supplemented with 1 µL/mL anti-clumping agent (Life Technologies). CHO-S WT and non-producing parental 10x KO cell lines were additionally supplemented with 8 mM L-glutamine. *rh*A1AT/*rh*C1INH producing clones were cultivated in L-glutamine-free medium at all times and passaged in medium

containing 25 µM MSX until the batch cultivation was initiated. Cell densities and viabilities were determined once per day using the NucleoCounter NC-250 Cell Counter (ChemoMetec) until the viability was <70%, at which point the culture was terminated. Supernatant samples with total secreted protein (secretome) from CHO-S WT and parental, non-producing 10x KO cell lines were taken five days after seeding and pooled within biological replicates. The volume for secretome samples was calculated to harbor 20 x 10^6 cells. For all shake flasks, additional supernatant samples were taken by centrifuging 1 mL of cell suspension for 5 minutes at 1000 g and storage of supernatant at -80°C until further analysis.

2.9 *rh*A1AT and *rh*C1INH purification

*rh*A1AT and *rh*C1INH were purified using CaptureSelect affinity resins (Thermo Fisher Scientific) according to the manufacturer's instructions. *rh*A1AT was further purified by size exclusion chromatography on a Superdex 200 increase 10/300GL column (GE Healthcare) equilibrated in PBS.

2.10 Titer assessment of *rh*A1AT/*rh*C1INH producing clones

*rh*A1AT and *rh*C1INH titers were determined using biolayer interferometry on an Octet RED96 (Pall, Menlo Park, CA, USA) as described previously for A1AT [39]. After hydration in PBS, streptavidin biosensors (18-5021, Fortebio, Pall) were functionalized with CaptureSelect biotin anti-A1AT conjugate or CaptureSelect biotin anti-C1INH conjugate (Thermo Fisher Scientific) at 5 µg/mL in PBS, and blocked in PBS containing 1 µg/mL biocytin (600 and 300 s incubation steps, respectively). Standards were prepared in spent CHO-S medium using plasma-derived A1AT (Athens Research & Technology) at 100, 50, 25, 12.5, 6.3, 3.1 and 1.6 µg/mL or C1INH (R&D systems) at 40, 20, 10, 5, 2.5, 1.25 and 0.625 µg/mL. Samples and standards were diluted two-fold and contained 0.1% BSA w/v, 0.1% tween-20 v/v, and 500 mM NaCl. When needed, samples were further diluted to fall within the range of the standard dilution series. After equilibration in spent CHO-S medium (120 s), samples and standards were measured for 300 s with a shaking speed of 1000 rpm at 30°C. Regeneration was performed with 50 mM TRIS, 2 M MgCl2, pH 7.5. Assays were performed in 96-well black microplates (Greiner Bio-One, Kremsmünster, Austria). Octet System Data Analysis 7.1 software was used to calculate binding rates and absolute A1AT and C1INH concentrations.

2.11 SDS-PAGE, isoelectric focusing and PNGase treatment

SDS-PAGE was performed on Novex 4-12% Tris-Glycine mini gels and isoelectric focusing (IEF) was performed on Novex pH 3-10 IEF gels (Thermo Fisher Scientific) as per the manufacturer's

instructions. Deglycosylation with PNGase F was performed according to the manufacturer's instructions (New England Biolabs, Ipswich, MA).

2.12 Activity assays

A1AT inhibitory activity was determined using the EnzChek Elastase Assay Kit (Molecular Probes, Eugene, OR) according to the manufacturer's instructions. In short, A1AT (8.0, 4.0, 2.0, 1.0, 0.5, 0.25, 0.13, and 0.06 µM) was incubated with purified active porcine pancreatic elastase and fluorescently labelled substrate (DQ-elastin). Measurement of fluorescence was performed after 45 min at room temperature (Excitation: 485 nm, slit width 9.0 nm; Emission: 530 nm, slit width 13.5 nm).

C1INH inhibitory activity was determined using the Technochrom C1INH Assay Kit (TechnoClone, Vienna, Austria). In short, plasma containing C1INH activity (120%, 60%, 30%) and samples (~0.25 µM) were incubated with substrate-buffer mixture for 3 min at room temperature, after which 50% acetic acid was added. Extinction was measured at 405 nm.

2.13 N-Glycan analysis

N-glycans were derivatized with GlycoWorks RapiFluor-MS N-Glycan Kit (Waters, Milford, MA) according to the manufacturer's instruction. Briefly; 12 µg purified protein or 12 µl of 10x concentrated (Amicon Ultra-15, Merck) secretome sample were used for each sample. Labeled N-Glycans were analyzed by LC-MS as described previously [33]. Separation gradient from 30% to 43% 50 mM ammonium formate buffer and MS were run in positive mode. Amount of N-Glycan was measured by integrating the peaks with Thermo Xcalibur software (Thermo Fisher Scientific, Waltham, MA) giving the normalized, relative amount of the glycans.

3 Results

3.1 Growth profile and N-glycan profile of clonal 10x KO cell lines

The aim of our study was to produce *rh*A1AT and *rh*C1INH in CHO cells with N-glycan profiles similar to human *pl*A1AT and *pl*C1INH. Our approach was to engineer the heterogeneous N-glycan profile of CHO-S WT cells towards a homogeneous A2G2S2 N-glycan structure, which is the predominant N-glycan on *pl*A1AT/*pl*C1INH. To this end, we generated out-of-frame insertions or deletions (indels) in eight glycosyltransferases (MGAT4A, MGAT4B, MGAT5, ST3GAL3, ST3GAL4, ST3GAL6, B3GNT2, FUT8) as well as in the genes *SPPL3* and *GLUL* (Suppl. Table S4) over four successive rounds of

multiplexed CRISPR/Cas9 gene editing. Two clones with indels in the targeted genes were subjected to growth analysis and N-glycan profiling.

Two clones (10x KO A and 10x KO B) with out-of-frame indels in all ten gene targets were obtained and both showed a pronounced increase in batch culture longevity when compared to the parental CHO-S WT cell line (Fig. 1A). CHO-S WT reached maximal viable cell density of ~6 x 10^6 cells/mL on day five and cell viability declined rapidly to <50% on day 6. In contrast, the 10x KO A and 10x KO B clones had cell viabilities >75% until day 10 of the batch cultivation and reached higher maximal viable cell density than CHO-S WT.

N-glycan analysis of the CHO-S WT secretome resulted in more than 25 annotated N-glycan structures (Fig. 1B) where the A2G2S2 structure, predominantly found on *pl*A1AT and *pl*C1INH, was not detected. The majority of CHO-S WT N-glycans contained core-fucosylation. The N-glycans produced by CHO-S WT cells appear diverse and comprise high-mannose structures as well as non-galactosylated, fully and partially sialylated di-, tri- and tetra-antennary structures (all with alpha-2,3-linked sialic acids). A2FG2S2 was found as the main N-glycan on total secreted proteins of CHO-S WT. In contrast, the N-glycan profiles of 10x KO A and 10x KO B are more homogeneous (Fig. 1B) with all structures lacking core-fucosylation. In addition, only relatively small amounts of CHO-specific alpha-2,3-linked sialylation were present. After disruption of the targeted genes, the proportion of A2G2 within N-glycan structures of total secreted proteins was increased from 3.5% (CHO-S WT) to 79% in both 10x KO clones (Fig. 1C). We concluded that the 10x KO A and B clones were suitable host cell lines in our effort to generate humanized N-glycans.

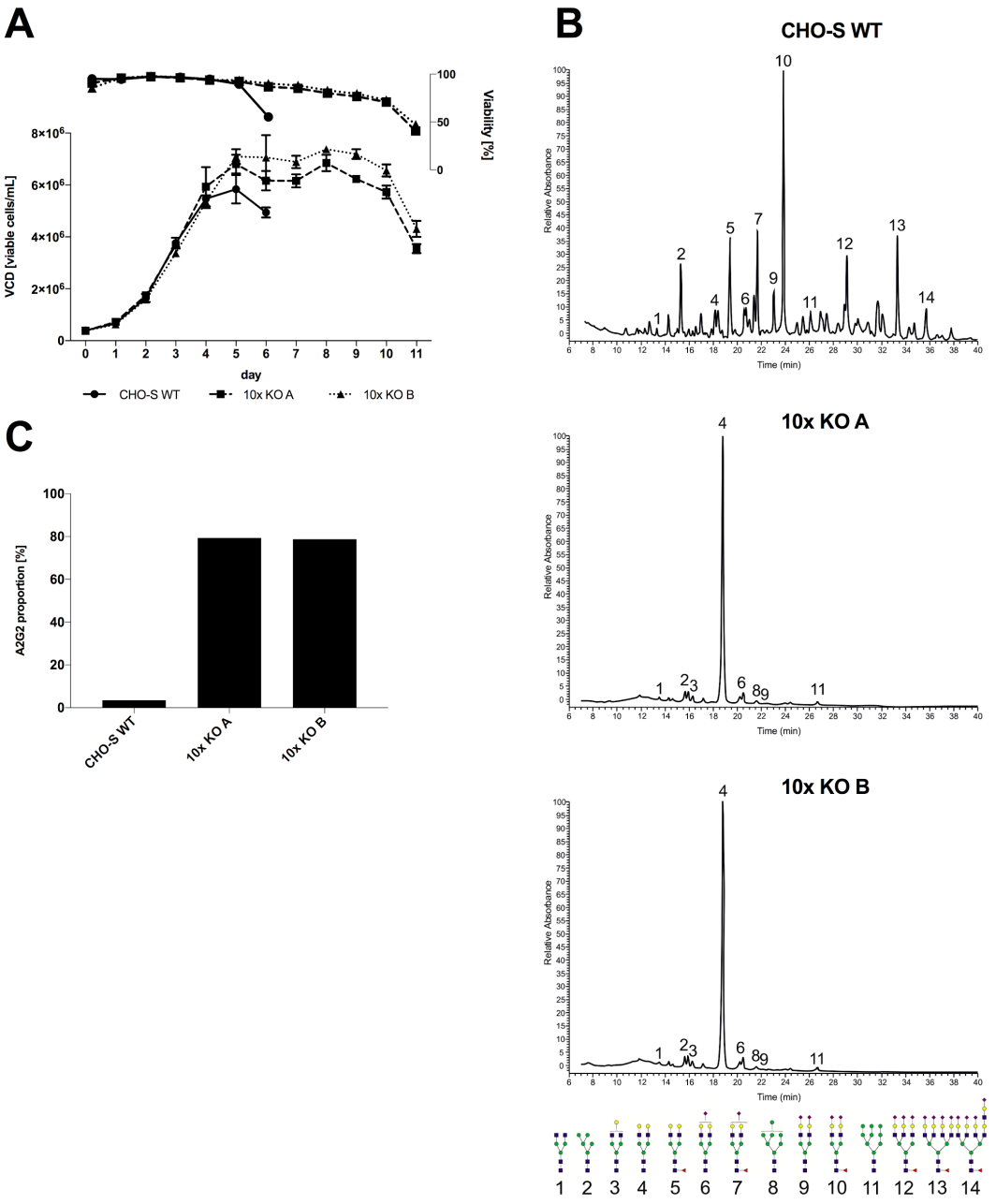


Fig. 1: Growth and N-glycan structure analysis of CHO-S WT and 10x KO cell lines. (A) Viable cell density (VCD) and viability of batch cultures of CHO-S WT and two clonal cell lines (10x KO A and 10x KO B) with indels in eight glycosyltransferases as well as GLUL and SPPL3. Error bars indicate the standard deviation of triplicate parallel cultures. (B) N-glycan profiling of total secreted proteins from CHO-S WT and the 10x KO A and 10x KO B clones. In the chromatogram, elution time indicated on the x-axis and y-axis represents signal intensity normalized to highest peak. (C) Proportion of non-fucosylated, biantennary N-glycans with terminal galactose (A2G2) in total secreted proteins from CHO-S WT and the 10x KO A and 10x KO B clones.

3.2 Introducing human-like sialylation in 10x KO cell lines

On the basis of A2G2 secretome N-glycan structures of clone 10x KO B, we aimed to develop clonal cell lines expressing St6gal1 and *rh*C1INH or St6gal1 and *rh*A1AT. We envisioned that such cell lines are capable to produce *rh*A1AT or *rh*C1INH with predominant A2G2S2 N-glycan structures as found on *pl*A1AT and *pl*C1INH. The functional GLUL-KO selection system was confirmed by MSX-dosage dependent recovery times of cell viabilities from transfected cell pools (Suppl. Fig. 3A). Passaging of the different transfection pools was performed until viability and doubling times were stable. We then conducted FACS-based single cell cloning with the 50 µM MSX-selected cells. During the expansion of the generated clones, only clones exhibiting predominant FITC-SNA staining and detectable levels of *rh*A1AT/*rh*C1INH in supernatants on coomassie-stained SDS-PAGE gels were selected (Suppl. Fig. 2). Based on these criteria, two *rh*A1AT (A1-1 and A1-2) and two *rh*C1INH (C1-1 and C1-2) producing clones were selected for further characterization.

SNA lectins are reported to bind predominantly to sialic acids of N-glycans linked to the galactose residue in a human-like alpha-2,6-sialylation. Analyzing FITC-SNA-stained CHO-S WT, we found relatively low levels of alpha-2,6-sialylation (Fig. 2A). To determine the proportion of cells with human-like sialylation, FITC-SNA stained CHO-S WT samples were used to gate between FITC-positive and FITC-negative cells (Suppl. Fig. 2A). Within the two 50 µM MSX-selected polyclonal cell lines, <30% of the cells were found to comprise alpha-2,6-linked sialic acids on N-glycans of cell surface proteins (Fig. 2B). 2B). In comparison, 82-90% of the cells in the populations of the selected four clones (A1-1, A1-2, C1-1 and C1-2) had the desired alpha-2,6-linked sialic acids on their N-glycans.

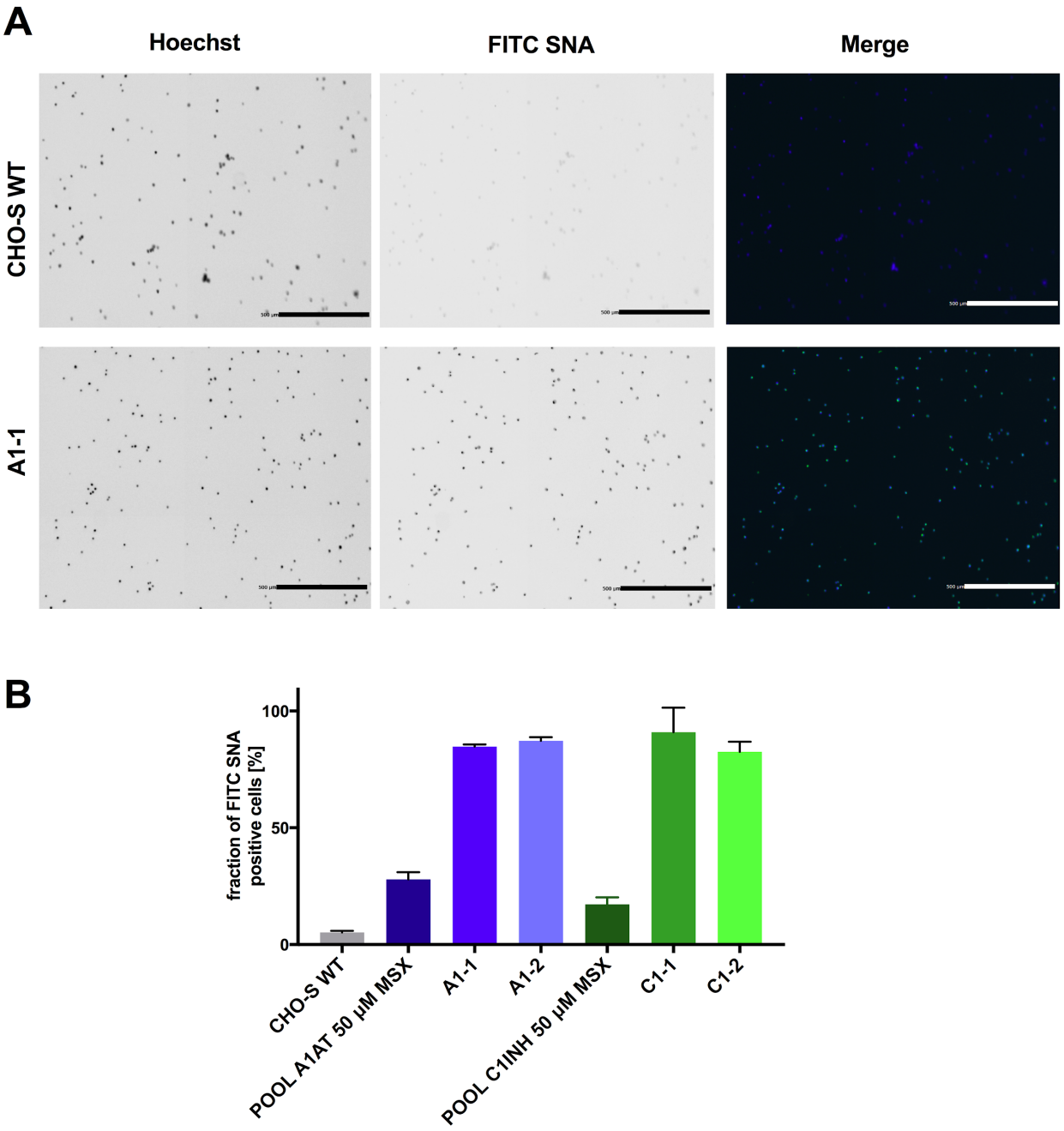


Fig. 2: FITC-SNA lectin staining of selected poly- and monoclonal cell lines. (A) Fluorescent images of CHO-S WT and A1-1 cell line. Cells were stained for alpha-2,6-sialic acid linkage with FITC-SNA (green) and for nuclei with Hoechst (blue). The bottom right corner bar displays a length of 500 µm (B) Comparison of FITC-SNA positive cells. FITC-SNA lectin staining of CHO-S WT, two 50 µM MSX polyclonal cell lines and four selected clones. Bars indicate the proportion of cells with positive FITC signal due to SNA lectin binding on alpha-2,6-linked sialic acids on the cell surface. Error bars represent standard deviation of three individual measurements per sample.

3.3 Extended culture longevity retained in *rh*A1AT and *rh*C1INH-producing clones

Both the two *rh*A1AT- and the two *rh*C1INH-producing clones showed the extended culture longevity as well as maximum VCD similar to the parental 10x KO B cell line (Fig. 3A). Viabilities of all clones were >75% until day 9 and clone C1-2 maintained cell viability >75% until day 11, similar to the non-producing parental 10x KO B.

In both *r*Ch1INH-producing clones, *rh*C1INH titers increased from day 0 until day 5 but then stagnated at ~40 µg/mL until the end of the batch cultures (Fig. 3B). In comparison, *rh*A1AT titers from clones A1-1 and A1-2 increased continuously to 123 µg/mL and 117 µg/mL , respectively. Despite increasing numbers of viable cells, the stagnation of C1INH titers at ~40 µg/mL for clones C1-1 and C1-2 in the second half of the batch cultures leads to the assumption that *rh*C1INH is unstable in the cell culture. By SDS-PAGE gel analysis of late phase supernatant samples we observed protein bands migrating just below *rh*C1INH (~70-100 kDa) which are not present in the non-producing parental clone (Suppl. Fig. 4). On the basis of titers and integral of viable cells (IVC) we determined the average specific productivity of the four clones during day 2 – 5 and day 6 - 9 (Fig. 3C). In the early phase, the specific productivity of *rh*A1AT for clones A1-1 and A1-2 was 5.8 and 4.0 pg/cell*day, respectively, decreasing to 2 – 3 pg/cell*day in the late phase (day 6 – 9). On the contrary, the two *rh*C1INH-producing clones expressed *rh*C1INH at ~5 pg/cell*day in the early phase, whereas their specific productivity in the late phase decreased to ~0 pg/cell*day. Overall, cell growth of all four clones was comparable to the parental cell line whereas product titers of A1AT-producing clones were increased compared to C1INH-producers.

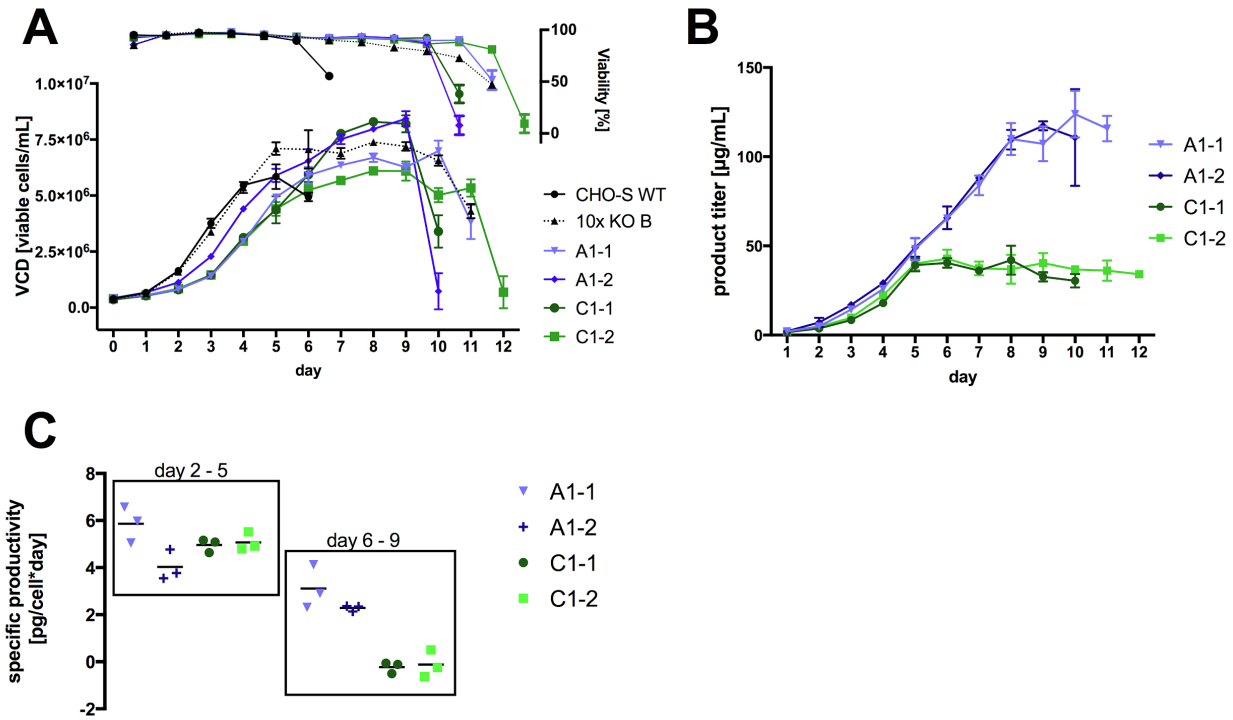


Fig. 3: Growth profiles, product titers and specific productivities of selected producing and non-producing clones. (A) Viable cell densities and cell viabilities of CHO-S WT, 10x KO B the *rh*A1AT- (A1-1 and A1-2) and *rh*C1INH- (C1-1 and C1-2) producing clonal cell lines measured in batch cultures. Error bars indicate range of duplicate parallel cultures. (B) *rh*A1AT and *rh*C1INH titer in supernatants during the batch culture experiment. Error bars indicate standard deviation of three individual measurements from two shake flasks per clone. (C) Specific productivities of the *rh*A1AT and *rh*C1INH-producing clonal cell lines in the batch culture experiment. Average specific productivity was calculated from day 2 – 5 and from day 6 – 9. Colored symbols represent average measured specific productivity for shake flask duplicates. Black lines shows the average specific productivity based on the three measurements of shake flask duplicates.

3.4 Activity and N-glycosylation profile of CHO-produced *rh*A1AT and *rh*C1INH are similar to plasma-derived products

We purified and characterized *rh*A1AT and *rh*C1INH to investigate the impact of our N-glycosylation engineering approach on product quality and protein activity. Therefore we compared the protein products produced in clonal cell lines derived from 10x KO B (*rh*A1AT and *rh*C1INH) to the CHO-S WT and plasma-derived counterparts (*pl*A1AT and *pl*C1INH).

SDS-PAGE gel analysis revealed that purified *rh*A1AT and *rh*C1INH produced in the four clones seem to have hydrodynamic volumes (molecular weight) similar to *pl*A1AT and *pl*C1INH without detectable impurities as seen in *pl*C1INH (Fig. 4A). *rh*A1AT and *rh*C1INH produced in CHO-S WT background did not co-migrate with *pl*A1AT and *pl*C1INH, respectively. However, after

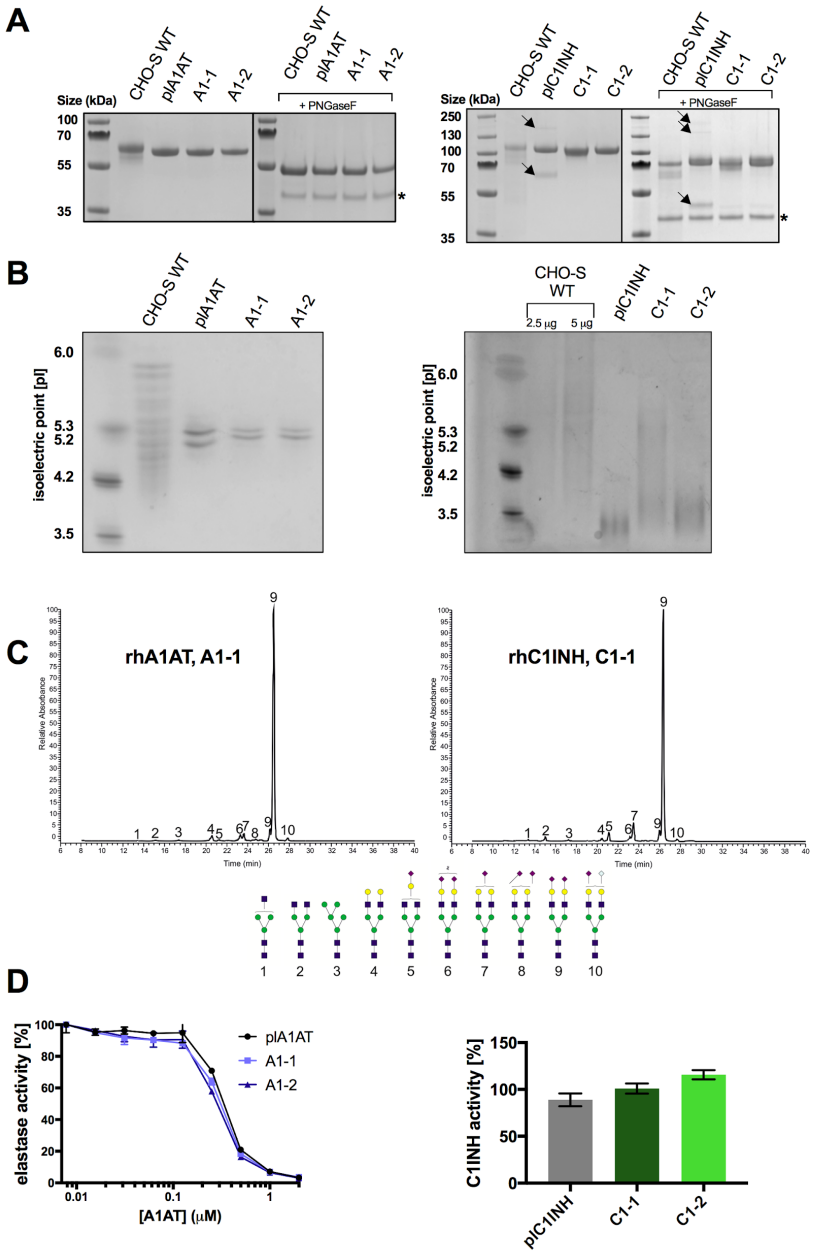


Fig. 4: Characterization of purified *rh*A1AT and *rh*C1INH. (A) SDS-PAGE gel analysis of commercially available Cinryze (*pl*C1INH) and Prolastin-C (*pl*A1AT) as well as *rh*A1AT and *rh*C1INH purified from polyclonal CHO-S WT or from monoclonal cell lines derived from 10x KO B. Removal of N-glycans by PNGaseF was performed where indicated. PNGaseF migrating as a ~40 kDa band is indicated with an asterisk and impurities of *pl*C1INH are indicated with arrows. (B) IEF gel analysis of same proteins as described for panel A. 2.5 µg purified protein was analysed per sample if not indicated otherwise. (C) N-glycan structures annotated from *rh*A1AT and *rh*C1INH produced in clones A1-1 and C1-1. (D) Left panel: *In vitro* assay measuring the inhibition of elastase activity at different concentrations of *pl*A1AT and *rh*A1AT purified from clones A1-1 and A1-2. Error bars indicate range of duplicate measurements. Maximum proteolytic activity of porcine elastase was set to 100%. Right panel: *In vitro* activity assessment of *pl*C1INH and *rh*C1INH purified from clones C1-1 and C1-2. As described in the assay, 1 IU/ml C1INH activity was set to 100%. Error bars indicate range of duplicate measurements.

deglycosylation with PNGaseF, all recombinantly produced proteins aligned with corresponding bands of *pl*A1AT and *pl*C1INH with the exception of *rh*C1INH produced in a CHO-S WT background displayed an additional protein band at ~65 kDa.

To further characterize the CHO-produced *rh*A1AT and *rh*C1INH, we performed IEF gel analysis (Fig. 4B). *rh*A1AT from clones A1-1 and A1-2 manifested in two bands with isoelectric points (pI) at pH 5.2 and 5.3 similar to *pl*A1AT. In contrast, *rh*A1AT produced in a CHO-S WT background displayed more than nine detectable isoforms with pI between pH 4 – 6.

IEF gel analysis of *rh*C1INH produced in a CHO-S WT background resulted in isoforms with pI ranging from pH ~4 – 6.5. A high degree of heterogeneity was also found in purified *rh*C1INH produced in clone C1-1. However, *rh*C1INH produced in clone C1-2 was less heterogeneous with pI at pH ~3.5 similar to *pl*C1INH.

In N-glycan analysis of purified *rh*A1AT and *rh*C1INH from CHO-S WT cells we detected a higher degree of heterogeneity compared to N-glycan structures on *rh*A1AT and *rh*C1INH from polyclonal 10x KO cell pools (Suppl. Fig. 3B). The polyclonal cell lines revealed two predominant sugar structures on both proteins (A2G2 and A2G2S2 N-glycans), whereas we could not detect the A2G2S2 structure on products from CHO-S WT. Moreover, the amount of predominant N-glycan structures on *rh*A1AT and *rh*C1INH was decreased from two (polyclonal pools) to one (monoclonal producers), identified as A2G2S2 N-glycan (Fig. 4C).

All four 10x KO-derived monoclonal cell lines produced *rh*A1AT and *rh*C1INH with higher proportion of A2G2S2 structures than *pl*A1AT and *pl*C1INH (Suppl. Fig. 2C and Suppl. Fig. 3C). The proportion of A2G2S2 in *rh*A1AT and *rh*C1INH was approximately 88 - 92% and 84%, respectively, and 82% for *pl*A1AT and 66% for *pl*C1INH .

Finally, we investigated the activity of purified *rh*A1AT and *rh*C1INH. *rh*A1AT activity was determined by its inhibitory function of elastase activity (Fig. 4D). Similar to *pl*A1AT, a decrease in elastase activity was detected at A1AT concentrations >0.1 µM for *rh*A1AT from clones A1-1 and A1-2. In addition, 50% of elastase inhibition was reached at ~0.3 µM A1AT for *pl*A1AT as well as *rh*A1AT. *In vitro* activity of purified *rh*C1INH produced by clones C1-1 and C1-2 was similar or higher compared to *pl*C1INH.

4 Discussion

We aimed to produce *rh*A1AT and *rh*C1INH in CHO-S with N-glycan profiles similar to *pl*A1AT and *pl*C1INH, which to our knowledge has not yet been achieved by recombinant expression. First, the heterogeneous N-glycan profile of CHO-S WT cells was changed to more homogeneous profiles in bespoke cell lines with predominant A2G2 N-glycan structures (Fig. 1B). Disrupting nine N-glycosylation-related genes increased the A2G2 proportion on total secreted protein from 3.5% in CHO-S WT-derived cells to ~80% in 10x KO cell lines (Fig. 1C). This supports the previously suggested strategy to decrease N-glycan branching and alpha-2,3-sialylation by disrupting MGAT4A, MGAT4B, MGAT5, ST3GAL3, ST3GAL4 and ST3GAL6 [34]. The impact of gene disruptions on cell culture performance was assessed in batch cultures. Interestingly, the monoclonal cell lines with disruption in ten gene targets showed enhanced growth characteristics compared to CHO-S WT cells (Fig. 1A). Overexpression of the GLUL gene has previously been found to decrease ammonia levels, which might explain improved CHO cell growth of the four characterized producer clones in L-glutamine-free medium [39]. However, the cause for the boosted cell growth of GLUL-lacking 10x KO cell lines in L-glutamine-supplemented medium (Fig. 1A) remains to be explored in further studies.

Since the disruption of the ten targets did not seem to interfere with cell culture performance, we performed co-expression of ST6GAL1 and *rh*A1AT or St6gal1 and *rh*C1INH in the 10x KO-derived clone B. After transfection, we observed a MSX-concentration dependent recovery of the transfected cell pools and successful killing of untransfected 10x KO B after 5 days of growth in L-glutamine-free medium similar to a previous study (Suppl. Fig. 3A) [40]. However, untransfected CHO-S WT cells were also able to recover up to the highest MSX-concentration of 50 µM, which is in accordance to previous work [36]. As shown in the killing curve of the untransfected 10x KO cells, the advantage of the GLUL-KO system here seems to be the elimination of untransfected cells.

Surprisingly, after selection at 50 µM MSX, the polyclonal cell lines did not show the desired predominant A2G2S2 glycosylation of purified *rh*A1AT and *rh*C1INH as we found incomplete sialylation on both proteins (Suppl. Fig. 3B). Similar lack of sialylation in the polyclonal cell lines was found after FITC-SNA lectin staining where only <30% of cells were identified to have alpha-2,6-sialylation (Fig. 2B). This might be due to incomplete vector integration into the genomic DNA or chromosome instability leading to a heterogeneous cell population with reduced stability of the integrated elements as reported earlier [41]. However, we were able to discard clones with incomplete sialylation by single cell cloning and screening for FITC-SNA lectin positive clones producing *rh*A1AT or *rh*C1INH.

Although the stagnating *rh*C1INH titer at day 5, the reported *rh*C1INH titer here is higher than previously reported production platforms with maximum titers of only 6 µg/mL in insect cells [29] and 30 µg/mL in *P. pastoris* [27] albeit the media we used was without any optimization for protein production. In comparison to *pl*C1INH (Cinryze), the *rh*C1INH from C1-1 and C1-2 did not show α1-antichymotrypsin, ceruloplasmin and Factor C3 impurities on SDS-PAGE gels [8] (Fig. 4A), implying that CHO-based cell platforms possibly can supply HAE-C1INH patients with higher purity than human plasma.

With titers between 300 – 400 µg/mL, human neuronal cell lines produce *rh*A1AT with higher titers than our clones A1-1 and A1-2. However, *rh*A1AT from these neuronal cells exhibits core-fucosylation, is not fully sialylated and therefore differs largely from *pl*A1AT N-glycosylation [24]. Similarly, earlier studies expressed *rh*A1AT in CHO with titers of up to 1.15 g/L, though differing from *pl*A1AT by revealing core-fucosylation and alpha-2,3-sialylation [25, 26, 42].

In contrast to the production platforms listed earlier, *rh*A1AT and *rh*C1INH produced in our 10x KO cell lines are not only exceeding sialylation levels of *pl*A1AT and *pl*C1INH (Suppl. Fig. 2C) but also reveal human-like alpha-2,6-sialylation instead of alpha-2,3-sialylation. Previous work reported only 2.6% A2G2S2 structures on *rh*A1AT expressed in CHO-K1, however with 2,3- sialic-acid linkage [26, 43].

Interestingly, the increased sialylation of *rh*A1AT from the two clones had no impact on *in vitro* activity (Fig. 4D). This is in accordance with previous work, which showed that A1AT activity is not linked to its N-glycosylation and CHO WT produced *rh*A1AT has similar activity to *pl*A1AT [10, 38]. Furthermore, differences of CHO-S WT- and 10x KO-derived *rh*A1AT were made visible using IEF gel analysis, where *rh*A1AT from the two clones revealed similar patterns to *pl*A1AT (Fig. 4B).

As presented in Fig. 4B, in contrast to *rh*A1AT, *rh*C1INH differed partially in IEF gel analysis profile from *pl*C1INH. Increased *rh*C1INH sialylation (Suppl. Fig. 2C) might lead to altered charge distribution and consequently cause changes in IEF gel patterns. For *rh*C1INH from clone C1-1, we observed a double-band in SDS-PAGE analysis although N-glycans were removed by PNGaseF treatment (Fig. 4A). This might indicate that O-glycan charge variants are responsible for the heterogeneity observed in IEF gel analysis and SDS-PAGE of *rh*C1INH from C1-1 (Fig. 4A and 4B) as described previously [44].

In summary, our work describes a strategy to successfully engineer the heterogeneous N-glycosylation profile of CHO-S WT cells towards the specific A2G2S2 N-glycan structure with the purpose of producing *rh*A1AT and *rh*C1INH with N-glycan profiles similar to human plasma-derived products. We used CRISPR/Cas9 to disrupt ten genes and then overexpressed *rh*A1AT or *rh*C1INH

on a ST6GAL1- and GLUL-encoding plasmid. After selection with MSX and single cell cloning, we identified clones expressing *rh*A1AT or *rh*C1INH with titers of up to 124 µg/mL and 42 µg/mL, respectively (Fig. 3B). Purified *rh*A1AT and *rh*C1INH were similar to the plasma-derived counterparts judged by SDS-PAGE analysis (Fig. 4A), degree and type of sialylation (Fig. 2B, Suppl. Fig. 2C, Suppl. Fig. 3C) and *in vitro* activity (Fig. 4D). Thus, the work presented shows the promise and potential of replacing cost-intensive and possibly unsafe plasma-derived augmentation therapy for AATD and C1INH-HAE patients by CHO- produced *rh*A1AT and *rh*C1INH. This strategy is in compliance with the Medical and Scientific Advisory Council (MASAC) recommendation of replacing plasma-derived products with recombinant products for the treatment of diseases [45].

Acknowledgement

The authors thank Sara Petersen Bjørn, Yuzhou Fan and Patrice Menard for valuable guidance and support. The authors thank Karen Katrine Brøndum, Nachon Charanyanonda Petersen, Karoline Schousboe Fremming and Zulfiya Sukhova for excellent technical assistance with the FACS, MiSeq library preparation and cell cultivation, Helle Munck Petersen for assistance with the protein purification, Anna Koza and Mads Valdemar Anderson for assistance with the MiSeq analysis and Tune Wulff for support with the N-glycan analysis. This work was supported by the Novo Nordisk Foundation (NNF10CC1016517). T.A., H.F.K. and M.R.A. are receiving funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No. 642663.

Author contributions

A. H. H., S. K., H. G. H., M. R. A., H. F. K., B. V., J. A., S. N. and T. A. designed the experiments. T. A. and J. A. constructed the knockout cell lines. S. N. designed and cloned the expression plasmids. T. A. generated and analyzed the final cell lines, performed protein expression and wrote the manuscript. S. K. performed protein purification and analysis. A. H. H. conducted N-glycosylation analysis. A. H. H., S. K., H. G. H., J. A., M. R. A., B. V., H. F. K. and G. M. L. guided the project, contributed to experimental design and commented and corrected the manuscript.

Competing interests

A patent based on this work has been filed. The authors on the patent are A. H. H., S. K., B. V., H. F. K. . The International Patent Application No. is EP17204071. The remaining authors declare no competing financial interests.

5 References

[1] Zhu, J., Mammalian cell protein expression for biopharmaceutical production. *Biotechnol. Adv.* 2012, *30*, 1158–1170.

[2] Dalziel, M., Crispin, M., Scanlan, C.N., Zitzmann, N., et al., Emerging Principles for the Therapeutic Exploitation of Glycosylation. *Science (80-.).* 2014, *343*, 1235681.

[3] Jennewein, M.F., Alter, G., The Immunoregulatory Roles of Antibody Glycosylation. *Trends Immunol.* 2017.

[4] Clerc, F., Reiding, K.R., Jansen, B.C., Kammeijer, G.S.M., et al., Human plasma protein N-glycosylation. *Glycoconj. J.* 2015, 1–35.

[5] Stoller, J.K., Aboussouan, L.S., α1-antitrypsin deficiency. *Lancet* 2005, *365*, 2225–2236.

[6] Feussner, A., Kalina, U., Hofmann, P., MacHnig, T., et al., Biochemical comparison of four commercially available C1 esterase inhibitor concentrates for treatment of hereditary angioedema. *Transfusion* 2014, *54*, 2566–2573.

[7] Mannucci, P.M., Castelli, R., Rumi, M.G., Agostoni, A., Reduction in transmission of hepatitis C after the introduction of a heat□treatment step in the production of C1□inhibitor concentrate. *Transfusion* 1995, *35*, 209–212.

[8] Filippi, F.D., Castelli, M., Cicardi, R., Soffredini, R., et al., Transmission of hepatitis G virus in patients with angioedema treated with steam-heated plasma concentrates of C1 inhibitor. *Transfusion* 1998, *38*, 307–311.

[9] Soucie, J.M., De Staercke, C., Monahan, P.E., Recht, M., et al., Evidence for the transmission of parvovirus B19 in patients with bleeding disorders treated with plasma-derived factor concentrates in the era of nucleic acid test screening. *Transfusion* 2013, *53*, 1217–1225.

[10] Karnaukhova, E., Ophir, Y., Golding, B., Recombinant human alpha-1 proteinase inhibitor: Towards therapeutic use. *Amino Acids* 2006, *30*, 317–332.

[11] Kwon, K.S., Yu, M.H., Effect of glycosylation on the stability of α1-antitrypsin toward urea denaturation and thermal deactivation. *Biochim. Biophys. Acta - Gen. Subj.* 1997, *1335*, 265–272.

[12] Stavenhagen, K., Kayili, H.M., Holst, S., Koeleman, C., et al., N- and O-glycosylation analysis of human C1-inhibitor reveals extensive mucin-type O-glycosylation. *Mol. Cell. Proteomics* 2017.

[13] Minta, J.O., The role of sialic acid in the functional activity and the hepatic clearance of C1-INH. *J. Immunol.* 1981, *126*, 245–9.

[14] Longhurst, H., Rhucin, a recombinant C1 inhibitor for the treatment of hereditary angioedema and cerebral ischemia. *Curr. Opin. Investig. Drugs* 2008, *9*, 310–323.

[15] van Doorn, M.B.A., Burggraaf, J., van Dam, T., Eerenberg, A., et al., A phase I study of recombinant human C1 inhibitor in asymptomatic patients with hereditary angioedema. *J. Allergy Clin. Immunol.* 2005, *116*, 876–83.

[16] Lewis, N.E., Liu, X., Li, Y., Nagarajan, H., et al., Genomic landscapes of Chinese hamster ovary cell lines as revealed by the Cricetulus griseus draft genome. *Nat. Biotechnol.* 2013, *31*, 759–65.

[17] Wright, G., Carver, A., Cottom, D., Reeves, D., et al., High level expression of active human alpha-1-antitrypsin in the milk of transgenic sheep. *Bio/Technology* 1991, *9*, 830–834.

[18] Ziomek, C.A., Commercialization of proteins produced in the mammary gland, in: *Theriogenology*, 1998, pp. 139–144.

[19] Moir, D.T., Dumais, D.R., Glycosylation and secretion of human alpha-1-antitrypsin by yeast. *Gene* 1987, *56*, 209–217.

[20] Khatami, M., Hosseini, S.N., Hasannia, S., Co-expression of alpha-1 antitrypsin with cytoplasmic domain of v-SNARE in Pichia pastoris: Preserving biological activity of alpha-1 antitrypsin. *Biotechnol. Appl. Biochem.* 2017.

[21] Jaberie, H., Naghibalhossaini, F., Recombinant production of native human α-1-antitrypsin protein in the liver HepG2 cells. *Biotechnol. Lett.* 2016, *38*, 1683–1690.

[22] Castilho, A., Windwarder, M., Gattinger, P., Mach, L., et al., Proteolytic and N-Glycan Processing of Human 1-Antitrypsin Expressed in Nicotiana benthamiana. *PLANT Physiol.* 2014, *166*, 1839–1851.

[23] Niklas, J., Priesnitz, C., Rose, T., Sandig, V., et al., Metabolism and metabolic burden by α1-antitrypsin production in human AGE1.HN cells. *Metab. Eng.* 2013, *16*, 103–114.

[24] Blanchard, V., Liu, X., Eigel, S., Kaup, M., et al., N-glycosylation and biological activity of recombinant human alpha1-antitrypsin expressed in a novel human neuronal cell line. *Biotechnol. Bioeng.* 2011, *108*, 2118–2128.

[25] Paterson, T., Innes, J., Moore, S., Approaches to maximizing stable expression of α1-antitrypsin in transformed CHO cells. *Appl. Microbiol. Biotechnol.* 1994, *40*, 691–698.

[26] Lee, K.J., Lee, S.M., Gil, J.Y., Kwon, O., et al., N-glycan analysis of human alpha-1-antitrypsin produced in Chinese hamster ovary cells. *Glycoconj. J.* 2013, *30*, 537–547.

[27] Bos, I.G.A., De Bruin, E.C., Karuntu, Y.A., Modderman, P.W., et al., Recombinant human C1-inhibitor produced in Pichia pastoris has the same inhibitory capacity as plasma C1-inhibitor. *Biochim. Biophys. Acta - Proteins Proteomics* 2003, *1648*, 75–83.

[28] Lamark, T., Ingebrigtsen, M., Bjørnstad, C., Melkko, T., et al., Expression of active human C1 inhibitor serpin domain in Escherichia coli. *Protein Expr. Purif.* 2001, *22*, 349–58.

[29] Wolff, M.W., Zhang, F., Roberg, J.J., Caldwell, E.E., et al., Expression of C1 esterase inhibitor by the baculovirus expression vector system: preparation, purification, and characterization. *Protein Expr. Purif.* 2001, *22*, 414–421.

[30] Frank, M.M., Recombinant and Plasma-Purified Human C1 Inhibitor for the Treatment of Hereditary Angioedema, in: *World Allergy Organization Journal*, 2010, pp. S29–S33.
[31] Wissing, S., Wölfel, J., Kewes, H., Niehus, C., et al., Expression of glycoproteins with excellent glycosylation profile and serum half-life in CAP-Go cells. *BMC Proc.* 2015, *9*, P12.
[32] Xu, X., Nagarajan, H., Lewis, N.E., Pan, S., et al., The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. *Nat. Biotechnol.* 2011, *29*, 735–741.
[33] Grav, L.M., Lee, J.S., Gerling, S., Kallehauge, T.B., et al., One-step generation of triple knockout CHO cell lines using CRISPR/Cas9 and fluorescent enrichment. *Biotechnol. J.* 2015, *10*, 1446–1456.
[34] Yang, Z., Wang, S., Halim, A., Schulz, M.A., et al., Engineered CHO cells for production of diverse, homogeneous glycoproteins. *Nat. Biotechnol.* 2015, *33*, 842–844.
[35] Ronda, C., Pedersen, L.E., Hansen, H.G., Kallehauge, T.B., et al., Accelerating genome editing in CHO cells using CRISPR Cas9 and CRISPy, a web-based target finding tool. *Biotechnol. Bioeng.* 2014, *111*, 1604–1616.
[36] Pristovšek, N., Hansen, H.G., Sergeeva, D., Borth, N., et al., Using Titer and Titer Normalized to Confluence Are Complementary Strategies for Obtaining Chinese Hamster Ovary Cell Lines with High Volumetric Productivity of Etanercept. *Biotechnol. J.* 2018, *13*.
[37] Lund, A.M., Kildegaard, H.F., Petersen, M.B.K., Rank, J., et al., A Versatile System for USER Cloning-Based Assembly of Expression Vectors for Mammalian Cell Engineering. *PLOS ONE* 2014, 9(5): e96693
[38] Hansen, H.G., Kildegaard, H.F., Lee, G.M., Kol, S., Case study on human alpha1-antitrypsin: Recombinant protein titers obtained by commercial ELISA kits are inaccurate. *Biotechnol. J.* 2016, *11*, 1648–1656.
[39] Noh, S.M., Park, J.H., Lim, M.S., Kim, J.W., et al., Reduction of ammonia and lactate through the coupling of glutamine synthetase selection and downregulation of lactate dehydrogenase-A in CHO cells. *Appl. Microbiol. Biotechnol.* 2017, *101*, 1035–1045.
[40] Noh, S.M., Shin, S., Lee, G.M., Comprehensive characterization of glutamine synthetase-mediated selection for the establishment of recombinant CHO cells producing monoclonal antibodies. *Sci. Rep.* 2018, *8*, 5361.
[41] Jun, S.C., Kim, M.S., Hong, H.J., Lee, G.M., Limitations to the development of humanized antibody producing chinese hamster ovary cells using glutamine synthetase-mediated gene amplification. *Biotechnol. Prog.* 2006, *22*, 770–780.
[42] Chin, C.L., Chin, H.K., Chin, C.S.H., Lai, E.T., et al., Engineering selection stringency on expression vector for the production of recombinant human alpha1-antitrypsin using Chinese Hamster ovary cells. *BMC Biotechnol.* 2015, 15, 44
[43] Lee, K.J., Lee, S.M., Gil, J.Y., Kwon, O., et al., N-glycan analysis of human α1-antitrypsin produced in Chinese hamster ovary cells. *Glycoconj. J.* 2013, 30(5):537-47
[44] Hansen, H.G., Nilsson, C.N., Lund, A.M., Kol, S., et al., Versatile microscale screening platform for improving recombinant protein productivity in Chinese hamster ovary cells. *Sci. Rep.* 2015, *5*, 18016.
[45] National Hemopilia Foundation, https://www.hemophilia.org/Researchers-Healthcare-Providers/Medical-and- Scientific-Advisory-Council-MASAC/MASAC-Recommendations/MASAC-Recommendations-Regarding-Standards-of-Service-for-Pharmacy-Providers-of-Clotting-Factor-Concentrates-for-Home-Use-to-Pat. 2014.

Chapter 5

Concluding remarks

Besides the well-established CHO-derived cell lines, other host platforms are on the rise to produce the increasing amount and variety of biopharmaceuticals. The cost-effective and time saving protocols for developing novel expression platforms is mostly driven by available genetic engineering tools as CRISPR/Cas9. However, with the existing history of approved biopharmaceuticals produced in CHO and the consolidated knowledge of bioprocessing and genetic modifications, CHO cells are superior to most other expression platforms. Due to recent and ongoing platform improvement efforts within industry but also several research groups[1] CHO cells will most likely continue their domination as the predominant working horse of the pharmaceutical industry. In order to provide novel CHO cell lines that enable the production of therapeutic proteins with homogeneous and/or human-like N-glycan structures, we disrupted up to ten target genes combined with the overexpression of a human glycosyltransferase. The possibility of multiplexing several gene targets simultaneously speeds up the rational design of improved CHO cell platforms significantly.

We explored that our protocol for simultaneous multiplexing of ten gene targets in a single round of transfection gave rise to undesired in-frame indels and therefore a inconvenient high number of clones to screen for cell lines with exclusively out-of-frame indels in all ten targets. To keep the screening efforts at a reasonable volume, we recommend to improve sgRNA efficacy, encode sgRNAs and Cas9 on the same plasmid and/or decrease the amount of targeted genes in each multiplexing round.

Based on the work of our explorative study described in Chapter 2, we reduced the maximum amount of simultaneously multiplexed gene targets to five targets per multiplexing round for the work presented in Chapter 3 and 4. Compared to repeated rounds of single target knockouts, the presented strategy allowed us to generate numerous cell lines in an efficient and time saving manner. Our multiplex protocols might also find application for other types of CHO engineering where several gene targets need to be disrupted for a desired phenotype, e.g. reducing toxic by-products or deleting secreted host cell proteins.

Unlike for the ten disrupted genes for humanized A1AT/C1INH production, certain combinations of disrupted target genes for decreased galactosylation were possibly associated with slightly decreased cell growth. This exemplifies that N-glycan engineering via gene disruption can impact cell characteristics beyond the biopharmaceutical N-glycan structure which could be disadvantageous from certain economic perspectives. Therefore one must also study whether the disrupted combination of gene targets for altered glycosylation also affects cell growth or the ability to express and secrete recombinant proteins properly. However, we proved the feasibility of tailoring certain N-glycan structures, which only occur little or are absent in the original host cell line. While genetic engineering represents a powerful tool to control N-glycosylation, researchers could combine our

protocols with medium and process design - the classical engineering approaches - to further explore strategies for designed N-glycosylation.

The protocols and results of this thesis support the *de novo* design of optimal N-glycosylation profiles for different therapeutic glycoproteins, presumably also for proteins or sugar residues not covered in our studies. Our engineering approach will aid the production of therapeutic proteins with improved properties and therapeutic potential. Furthermore, we expand the possible scope of biopharmaceuticals by enabling the recombinant production of therapeutic proteins which require human-identical N-glycosylation.

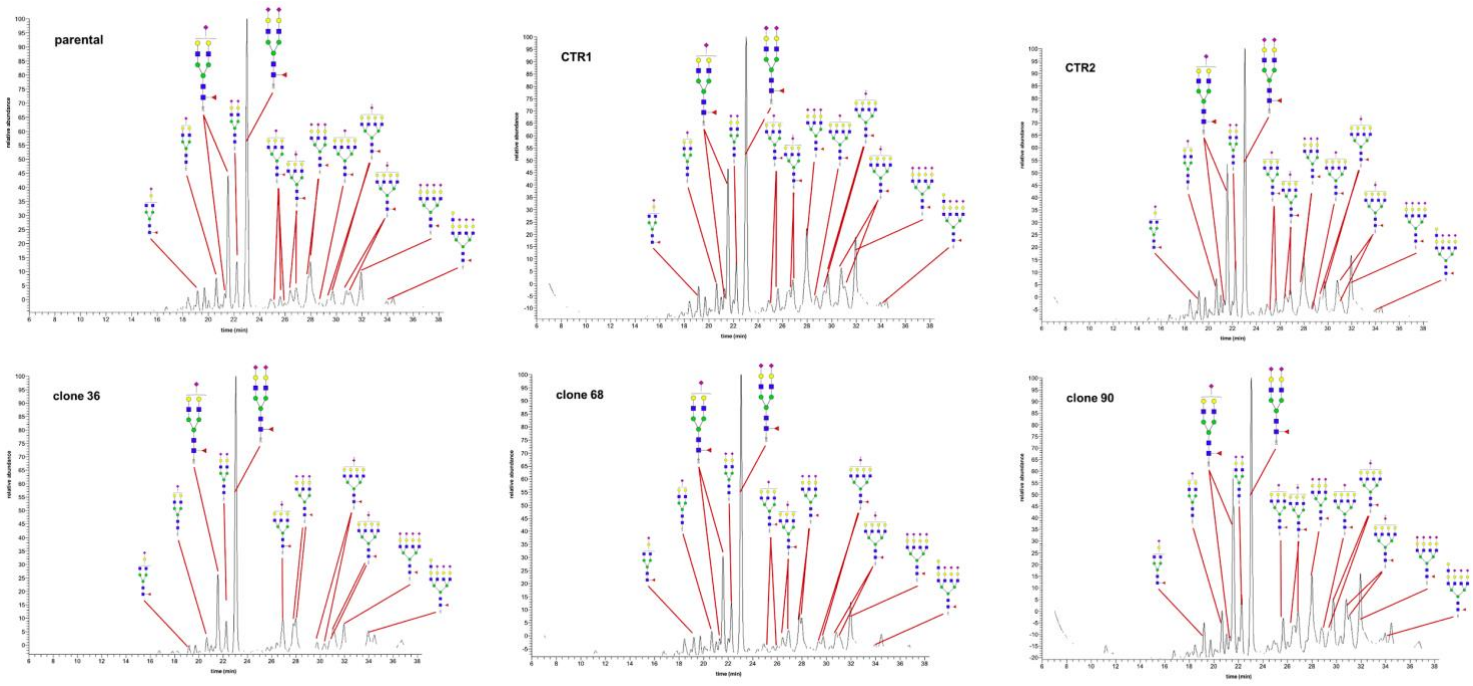
1. Fischer, S., Handrick, R. & Otte, K. The art of CHO cell engineering: A comprehensive retrospect and future perspectives. *Biotechnol. Adv.* **33,** 1878–1896 (2015).

Supplementary Material for

Deca CHO KO: exploring the limitations of CRISPR/Cas9 multiplexing in CHO cells

Thomas Amann[1], Anders Holmgaard Hansen[1], Gyun Min Lee[1,2], Mikael Rørdam Andersen[3], Helene Faustrup Kildegaard[1]

[1]The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark
[2]Department of Biological Sciences, KAIST, Daejeon, Republic of Korea
[3]Department of Biotechnology and Biomedicine, Technical University of Denmark, Kgs. Lyngby, Denmark



Supplementary Figure S1: Total secreted protein N-glycan analysis. Analysis of released N-glycans from total secreted proteins of the parental cell line, two control cell lines and three clones with indels in some of the targeted genes.

Supplementary Table S1: sgRNA target sequences. The bases in red mark the PAM site.

Target gene	Target sequence (5' → 3')
BAX	GCTGATGGCAACTTCAACTGGG
BAK1	GGAAGCCGGTCAAACCACGTGG
GLUL	GGCCCAGGGAAGCCATCGGAGG
SPPL3	AGAGAGACGGACGCTCCAATGG
B4GALT1	TATCCCATTTCGCAACCGGCGG
B4GALT2	GGAGCACCACCTACGCTATGG
B4GALT3	TGATGCTCGCGGGCACGATGG
B4GALT4	TGGAGGCCGGTATCACCCTGG
B4GALT5	GAGAGTGACCGCAACTACTAGG
TSTA3	GAGGTGGTCGCAGATGGCGCGG

Supplementary Table S2: Oligos for sgRNA expression vector cloning.

	Oligo sequence (5' → 3')
gRNA_BAX_1345650_fwd	GGAAAGGACGAAACACCGCTGATGGCAACTTCAACTGGTTTTAGAGCTAGAAAT
BAX_1345650_rev	CTAAAACCAGTTGAAGTTGCCATCAGCGGTGTTTCGTCCTTTCCACAAGATAT
gRNA_BAK1_1544257_fwd	GGAAAGGACGAAACACCGGAAGCCGGTCAAACCACGTGTTTTAGAGCTAGAAAT
BAK1_1544257_rev	CTAAAACACGTGGTTTGACCGGCTTCCGGTGTTTCGTCCTTTCCACAAGATAT
gRNA_GLUL_941540_fwd	GGAAAGGACGAAACACCGGCCCAGGGAAGCCATCGGAGTTTTAGAGCTAGAAAT
GLUL_941540_rev	CTAAAACTCCGATGGCTTCCCTGGGCCGGTGTTTCGTCCTTTCCACAAGATAT
gRNA_Sppl3_NW_003613978.1_213040_fwd	GGAAAGGACGAAACACCAGAGAGACGGACGCTCCAATGTTTTAGAGCTAGAAAT
gRNA_Sppl3_NW_003613978.1_213040_rev	CTAAAACATTGGAGCGTCCGTCTCTCTGGTGTTTCGTCCTTTCCACAAGATAT
gRNA_B4galt1_NW_003615120.1_283507_fwd	GGAAAGGACGAAACACCTATCCCATTTCGCAACCGGCGTTTTAGAGCTAGAAAT
gRNA_B4galt1_NW_003615120.1_283507_rev	CTAAAACGCCGGTTGCGAAATGGGATAGGTGTTTCGTCCTTTCCACAAGATAT
gRNA_B4galt2_NW_003613906.1_163085_fwd	GGAAAGGACGAAACACCGGGAGCACCACCTACGCTATGTTTTAGAGCTAGAAAT
gRNA_B4galt2_NW_003613906.1_163085_rev	CTAAAACATAGCGTAGGTGGTGCTCCCGGTGTTTCGTCCTTTCCACAAGATAT
gRNA_B4galt3_NW_003614301.1_233251_fwd	GGAAAGGACGAAACACCGTGATGCTCGCGGGCACGATGTTTTAGAGCTAGAAAT
gRNA_B4galt3_NW_003614301.1_233251_rev	CTAAAACATCGTGCCCGCGAGCATCACGGTGTTTCGTCCTTTCCACAAGATAT
gRNA_B4galt4_NW_003614660.1+444821_fwd	GGAAAGGACGAAACACCGGAGGCCGGTATCACCCTGGTTTTAGAGCTAGAAAT
gRNA_B4galt4_NW_003614660.1+444767_rev	CTAAAACAAGTGAGGTCCGGCTCAAAGGTGTTTCGTCCTTTCCACAAGATAT
gRNA_B4galt5_751697_fwd	GGAAAGGACGAAACACCGAGAGTGACCGCAACTACTAGTTTTAGAGCTAGAAAT
gRNA_B4galt5_751697_rev	CTAAAATAGTAGTTGCGGTCACTCTCCGGTGTTTCGTCCTTTCCACAAGATAT
Tsta3_NW_003617373.1-6731_gRNAfwd	GGAAAGGACGAAACACCGAGGTGGTCGCAGATGGCGCGTTTTAGAGCTAGAAAT
Tsta3_NW_003617373.1-6731_gRNArev	CTAAAACGCGCCATCTGCGACCACCTCGGTGTTTCGTCCTTTCCACAAGATAT

Supplementary Table S3: Cas9_2A_GFP and sgRNA plasmid ratio for transfection.

plasmid	µg for transfection
Cas9_2A_GFP	1.88
sgRNA BAX	0.19
sgRNA BAK1	0.19
sgRNA GLUL	0.19
sgRNA SPPL3	0.19
sgRNA B4GALT1	0.19
sgRNA B4GALT2	0.19
sgRNA B4GALT3	0.19
sgRNA B4GALT4	0.19
sgRNA B4GALT5	0.19
sgRNA TSTA3	0.19

Supplementary Table S4: Primer list for deep sequencing (MiSeq). The primers contain overhang sequences compatible with Illumina Nextera XT indexing (forward primer overhang: TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG, reverse primer overhang: GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG).

	Oligo sequence (5' → 3')
MiSeq_BAX_1345650_fwd	**TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTGTGGATACTAACTCCCCACG**
MiSeq_BAX_1345650_rev	**GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTCCCTGAACCTCACTACCCC**
MiSeq_BAK1_1544257_fwd	**TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCAAGGTGGGCTCTCCGTGAT**
MiSeq_BAK1_1544257_rev	**GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCGATGCAATGGTGCAGTATGAT**
MiSeq_GLUL_941540_fwd	**TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAAGACACACGTGTAAACGGA**
MiSeq_GLUL_941540_rev	**GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTCCCTGCGATAGGCTTTGTC**
MiSeq_Sppl3_NW_003613978.1_213040_fwd	**TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCGTGGAGTAACTTACCTGCTGT**
MiSeq_Sppl3_NW_003613978.1_213040_rev	**GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAAGTGGTGAGTGTGTCCTGT**
MiSeq_B4galt1_NW_003615120.1_283507_fwd	**TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGACGCATTTGTGTACCCGAGT**
MiSeq_B4galt1_NW_003615120.1_283507_rev	**GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAGATGGGCGGTCGTTATTCC**
MiSeq_B4galt2_NW_003613906.1_163085_fwd	**TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAGCAGACACATCCACAGGTG**
MiSeq_B4galt2_NW_003613906.1_163085_rev	**GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTCCTGGGAGGCCGTTATACA**
MiSeq_B4galt3_NW_003614301.1_233251_fwd	**TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTGGATTTCTGTATGGGGCCA**
MiSeq_B4galt3_NW_003614301.1_233251_rev	**GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGTGGGTCCTGTGTCGGTATC**
MiSeq_B4galt4_NW_003614660.1+444821_fwd	**TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAGCCCCTCTTTCGACATGTG**
MiSeq_B4galt4_NW_003614660.1+444821_rev	**GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTGCCATAGTCCAGTTGCTGC**
Tsta3_NW_003617373.16731_MiSeqfwd	**TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGCTCAGGAACAGACCCAACA**
Tsta3_NW_003617373.16731_MiSeqrev	**GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGTGCAGCAACAATGGGTGAG**

Supplementary Table S5: Deep sequencing results of single cell derived clones. Detailed sequencing analysis results for the targeted genes of generated clones. Indels are marked in green and in-frame indels highlighted in red.

Clone Nr	Target gene	Reads	WT	Indel	% Indel	peak 1	peak 2	peak 3	peak 4
wt	GS	14,381	14,326	55	0.38%				
wt	Tsta3	2,869	2,850	19	0.66%				
wt	B4galt4	4,878	4,866	12	0.25%				
wt	Sppl3	13,907	13,845	62	0.45%				
wt	Bax1	5,869	5,822	47	0.80%				
wt	B4galt1	2,281	2,270	11	0.48%				
wt	B4galt3	6,389	6,361	28	0.44%				
wt	B4galt2	8,016	7,995	21	0.26%				
wt	b4galt5	7,026	6,998	28	0.40%				
wt	Bak1	13,944	13,840	104	0.75%				
2	GS	13,382	13,330	52	0.39%				
2	Tsta3	1,102	1,095	7	0.64%				
2	B4galt4	5,153	5,134	19	0.37%				
2	Sppl3	8,865	8,812	53	0.60%				
2	Bax1	4,021	3,973	48	1.19%				
2	B4galt3	4,916	4,892	24	0.49%				
2	B4galt2	5,957	5,943	14	0.24%				
2	b4galt5	5,651	5,624	27	0.48%				
2	Bak1	10,868	10,807	61	0.56%				
3	GS	15,272	15,218	54	0.35%				
3	Tsta3	300	296	4	1.33%				
3	B4galt4	6,389	6,360	29	0.45%				
3	Sppl3	12,656	12,592	64	0.51%				
3	Bax1	6,277	6,225	52	0.83%				
3	B4galt1	5,396	5,379	17	0.32%				
3	B4galt3	5,514	5,489	25	0.45%				
3	B4galt2	7,772	7,761	11	0.14%				
3	b4galt5	6,466	6,438	28	0.43%				
3	Bak1	12,910	12,811	99	0.77%				
4	GS	12,952	12,901	51	0.39%				
4	Tsta3	912	907	5	0.55%				
4	B4galt4	6,056	6,027	29	0.48%				
4	Sppl3	9,600	9,555	45	0.47%				
4	Bax1	5,458	5,417	41	0.75%				
4	B4galt1	115	115	0	0.00%				
4	B4galt3	5,113	5,086	27	0.53%				
4	B4galt2	6,631	6,610	21	0.32%				
4	b4galt5	5,758	5,728	30	0.52%				
4	Bak1	10,553	10,488	65	0.62%				
5	GS	12,179	12,148	31	0.25%				
5	Tsta3	1,912	1,832	80	4.18%				
5	B4galt4	5,915	2,679	3,236	54.71%	-1	2		
5	Sppl3	10,541	10,495	46	0.44%				
5	Bax1	5,620	369	5,251	93.43%	-2	1		
5	B4galt1	2,062	2,055	7	0.34%				
5	B4galt3	4,624	4,574	50	1.08%				
5	B4galt2	6,024	5,992	32	0.53%				
5	b4galt5	4,994	4,962	32	0.64%				
5	Bak1	10,809	10,702	107	0.99%				
6	GS	17,463	17,399	64	0.37%				
6	Tsta3	1,568	1,560	8	0.51%				
6	B4galt4	8,599	8,518	81	0.94%				
6	Sppl3	15,560	15,482	78	0.50%				
6	Bax1	7,641	7,590	51	0.67%				
6	B4galt3	6,668	6,647	21	0.31%				
6	B4galt2	9,895	9,868	27	0.27%				
6	b4galt5	7,755	7,709	46	0.59%				
6	Bak1	14,872	14,755	117	0.79%				

7	GS	15,698	15,649	49	0.31%				
7	Tsta3	1,556	403	1,153	74.10%	-6 (inframe)	-2	-1	
7	B4galt4	7,840	1,783	6,057	77.26%	-1	1		
7	Sppl3	13,828	13,752	76	0.55%				
7	Bax1	7,624	25	7,599	99.67%	-3 (inframe)	1		
7	B4galt1	3,997	3,986	11	0.28%				
7	B4galt3	6,046	6,022	24	0.40%				
7	B4galt2	8,430	8,407	23	0.27%				
7	b4galt5	6,829	6,801	28	0.41%				
7	Bak1	13,426	10,931	2,495	18.58%	1			
8	GS	14,382	14,324	58	0.40%				
8	Tsta3	1,571	1,564	7	0.45%				
8	B4galt4	7,102	7,083	19	0.27%				
8	Sppl3	12,378	12,317	61	0.49%				
8	Bax1	6,696	6,651	45	0.67%				
8	B4galt1	4,302	4,290	12	0.28%				
8	B4galt3	5,943	5,924	19	0.32%				
8	B4galt2	7,774	7,741	33	0.42%				
8	b4galt5	6,412	6,376	36	0.56%				
8	Bak1	11,920	11,830	90	0.76%				
9	GS	13,479	13,432	47	0.35%				
9	Tsta3	926	923	3	0.32%				
9	B4galt4	6,584	6,552	32	0.49%				
9	Sppl3	11,874	11,823	51	0.43%				
9	Bax1	6,151	6,106	45	0.73%				
9	B4galt1	4,258	4,240	18	0.42%				
9	B4galt3	5,745	5,715	30	0.52%				
9	B4galt2	6,790	6,775	15	0.22%				
9	b4galt5	5,765	5,735	30	0.52%				
9	Bak1	11,855	11,766	89	0.75%				
10	GS	11,720	11,675	45	0.38%				
10	Tsta3	645	637	8	1.24%				
10	B4galt4	5,361	5,341	20	0.37%				
10	Sppl3	9,488	9,451	37	0.39%				
10	Bax1	5,714	5,663	51	0.89%				
10	B4galt1	3,516	3,497	19	0.54%				
10	B4galt3	4,333	4,320	13	0.30%				
10	B4galt2	5,873	5,857	16	0.27%				
10	b4galt5	5,226	5,196	30	0.57%				
10	Bak1	10,027	9,957	70	0.70%				
11	GS	17,815	17,741	74	0.42%				
11	Tsta3	1,801	1,790	11	0.61%				
11	B4galt4	9,988	9,943	45	0.45%				
11	Sppl3	14,198	14,122	76	0.54%				
11	Bax1	8,363	8,290	73	0.87%				
11	B4galt1	5,485	5,458	27	0.49%				
11	B4galt3	7,956	7,930	26	0.33%				
11	B4galt2	8,620	8,598	22	0.26%				
11	b4galt5	9,182	9,131	51	0.56%				
11	Bak1	15,905	15,795	110	0.69%				
12	GS	14,901	14,837	64	0.43%				
12	Tsta3	841	835	6	0.71%				
12	B4galt4	8,303	8,234	69	0.83%				
12	Sppl3	13,823	13,765	58	0.42%				
12	Bax1	7,938	7,882	56	0.71%				
12	B4galt1	2,742	2,736	6	0.22%				
12	B4galt3	6,552	6,521	31	0.47%				
12	B4galt2	8,342	8,318	24	0.29%				
12	b4galt5	7,686	7,658	28	0.36%				

12	Bak1	13,079	12,998	81	0.62%				
13	GS	15,854	15,801	53	0.33%				
13	Tsta3	1,708	1,697	11	0.64%				
13	B4galt4	7,091	7,069	22	0.31%				
13	Sppl3	7,997	7,953	44	0.55%				
13	Bax1	7,848	7,786	62	0.79%				
13	B4galt1	4,809	4,787	22	0.46%				
13	B4galt3	5,057	5,028	29	0.57%				
13	B4galt2	7,184	7,171	13	0.18%				
13	b4galt5	7,648	7,622	26	0.34%				
13	Bak1	13,180	13,096	84	0.64%				
14	GS	16,495	16,443	52	0.32%				
14	Tsta3	1,099	3	1,096	99.73%	-1			
14	B4galt4	7,568	14	7,554	99.82%	-1			
14	Sppl3	11,133	11,082	51	0.46%				
14	Bax1	5,172	7	5,165	99.86%	4			
14	B4galt1	4,990	4,972	18	0.36%				
14	B4galt3	6,085	6,061	24	0.39%				
14	B4galt2	7,274	7,099	175	2.41%				
14	b4galt5	7,584	7,541	43	0.57%				
14	Bak1	13,175	7,330	5,845	44.36%	-1	8		
15	GS	13,521	13,472	49	0.36%				
15	Tsta3	1,405	1,395	10	0.71%				
15	B4galt4	6,230	6,190	40	0.64%				
15	Sppl3	10,549	10,497	52	0.49%				
15	Bax1	6,590	6,532	58	0.88%				
15	B4galt1	6,753	6,724	29	0.43%				
15	B4galt3	5,582	5,557	25	0.45%				
15	B4galt2	6,952	6,935	17	0.24%				
15	b4galt5	5,742	5,715	27	0.47%				
15	Bak1	11,739	11,645	94	0.80%				
16	GS	13,713	13,670	43	0.31%				
16	Tsta3	1,299	5	1,294	99.62%	-6 (inframe)	-2	-1	
16	B4galt4	5,354	11	5,343	99.79%	-1			
16	Sppl3	11,545	6,783	4,762	41.25%	-17			
16	Bax1	4,393	35	4,358	99.20%	1			
16	B4galt1	6,429	6,399	30	0.47%				
16	B4galt3	4,227	2,227	2,000	47.31%	1			
16	B4galt2	5,485	2,614	2,871	52.34%	-12 (inframe)	1		
16	b4galt5	5,109	2,733	2,376	46.51%	-16			
16	Bak1	7,475	56	7,419	99.25%	-1			
17	GS	17,040	16,983	57	0.33%				
17	Tsta3	2,102	2,086	16	0.76%				
17	B4galt4	7,388	7,352	36	0.49%				
17	Sppl3	11,674	11,615	59	0.51%				
17	Bax1	6,967	6,909	58	0.83%				
17	B4galt1	7,629	7,602	27	0.35%				
17	B4galt3	6,316	6,285	31	0.49%				
17	B4galt2	7,220	7,190	30	0.42%				
17	b4galt5	7,227	7,188	39	0.54%				
17	Bak1	14,965	14,863	102	0.68%				
18	GS	16,852	16,785	67	0.40%				
18	Tsta3	2,640	2,625	15	0.57%				
18	B4galt4	8,217	8,183	34	0.41%				
18	Sppl3	15,733	15,620	113	0.72%				
18	Bax1	6,203	6,159	44	0.71%				
18	B4galt1	8,116	8,084	32	0.39%				
18	B4galt3	6,496	6,452	44	0.68%				
18	B4galt2	7,608	7,587	21	0.28%				
18	b4galt5	7,701	7,665	36	0.47%				
18	Bak1	13,817	13,735	82	0.59%				

19	GS	12,630	12,591	39	0.31%				
19	Tsta3	1,354	1,349	5	0.37%				
19	B4galt4	6,053	4,037	2,016	33.31%	-18 (inframe)			
19	Sppl3	9,097	9,053	44	0.48%				
19	Bax1	5,441	5,397	44	0.81%				
19	B4galt1	4,933	4,917	16	0.32%				
19	B4galt3	4,503	4,483	20	0.44%				
19	B4galt2	5,920	5,908	12	0.20%				
19	b4galt5	5,509	5,480	29	0.53%				
19	Bak1	10,345	10,276	69	0.67%				
20	GS	13,643	13,587	56	0.41%				
20	Tsta3	1,340	1,330	10	0.75%				
20	B4galt4	6,598	6,545	53	0.80%				
20	Sppl3	10,384	10,325	59	0.57%				
20	Bax1	5,799	5,765	34	0.59%				
20	B4galt1	6,466	6,443	23	0.36%				
20	B4galt3	5,143	5,128	15	0.29%				
20	B4galt2	6,482	6,469	13	0.20%				
20	b4galt5	5,193	5,162	31	0.60%				
20	Bak1	10,280	10,212	68	0.66%				
21	GS	13,779	13,717	62	0.45%				
21	Tsta3	1,172	1,166	6	0.51%				
21	B4galt4	6,375	6,343	32	0.50%				
21	Sppl3	8,618	8,588	30	0.35%				
21	Bax1	5,824	5,779	45	0.77%				
21	B4galt1	6,405	6,387	18	0.28%				
21	B4galt3	5,181	5,159	22	0.42%				
21	B4galt2	6,925	6,895	30	0.43%				
21	b4galt5	6,256	6,228	28	0.45%				
21	Bak1	11,195	11,124	71	0.63%				
22	GS	11,355	11,313	42	0.37%				
22	Tsta3	544	240	304	55.88%	-6 (inframe)	-2		
22	B4galt4	5,364	5,321	43	0.80%				
22	Sppl3	7,670	7,633	37	0.48%				
22	Bax1	5,118	2,685	2,433	47.54%	-2			
22	B4galt1	5,667	5,645	22	0.39%				
22	B4galt3	3,925	3,911	14	0.36%				
22	B4galt2	5,391	5,378	13	0.24%				
22	b4galt5	5,996	5,960	36	0.60%				
22	Bak1	9,805	9,148	657	6.70%	-21 (inframe)			
23	GS	11,826	11,774	52	0.44%				
23	Tsta3	1,133	1,125	8	0.71%				
23	B4galt4	5,332	5,305	27	0.51%				
23	Sppl3	9,982	9,947	35	0.35%				
23	Bax1	4,520	40	4,480	99.12%	-1	1		
23	B4galt1	5,445	5,428	17	0.31%				
23	B4galt3	4,404	4,390	14	0.32%				

23	B4galt2	5,335	5,324	11	0.21%				
23	b4galt5	4,634	3,721	913	19.70%	1			
23	Bak1	9,463	6,813	2,650	28.00%	-14	-12 (inframe)		
24	GS	19,088	19,031	57	0.30%				
24	Tsta3	199	198	1	0.50%				
24	B4galt4	8,537	8,503	34	0.40%				
24	Sppl3	14,999	14,915	84	0.56%				
24	Bax1	8,801	8,733	68	0.77%				
24	B4galt1	7,768	7,745	23	0.30%				
24	B4galt3	7,103	7,073	30	0.42%				
24	B4galt2	9,925	4,979	4,946	49.83%	-48 (inframe)	-18 (inframe)		
24	b4galt5	8,472	8,423	49	0.58%				
24	Bak1	15,165	15,005	160	1.06%				
25	GS	11,442	11,389	53	0.46%				
25	Tsta3	936	933	3	0.32%				
25	B4galt4	7,261	7,228	33	0.45%				
25	Sppl3	12,660	12,599	61	0.48%				
25	Bax1	7,843	7,780	63	0.80%				
25	B4galt3	4,085	4,064	21	0.51%				
25	B4galt2	8,335	8,309	26	0.31%				
25	b4galt5	7,371	7,323	48	0.65%				
25	Bak1	12,022	11,938	84	0.70%				
26	GS	12,591	12,543	48	0.38%				
26	Tsta3	1,917	1,904	13	0.68%				
26	B4galt4	6,409	6,380	29	0.45%				
26	Sppl3	13,094	12,995	99	0.76%				
26	Bax1	3,838	3,812	26	0.68%				
26	B4galt1	4,391	4,376	15	0.34%				
26	B4galt3	5,223	5,202	21	0.40%				
26	B4galt2	6,971	6,959	12	0.17%				
26	b4galt5	5,637	5,614	23	0.41%				
26	Bak1	11,953	11,850	103	0.86%				
27	GS	12,444	12,407	37	0.30%				
27	Tsta3	1,378	1,373	5	0.36%				
27	B4galt4	7,570	7,539	31	0.41%				
27	Sppl3	12,923	12,850	73	0.56%				
27	Bax1	5,991	5,942	49	0.82%				
27	B4galt1	3,743	3,727	16	0.43%				
27	B4galt3	6,057	6,032	25	0.41%				
27	B4galt2	7,876	7,855	21	0.27%				
27	b4galt5	7,010	6,985	25	0.36%				
27	Bak1	12,022	11,936	86	0.72%				
28	GS	12,418	12,369	49	0.39%				
28	Tsta3	2,017	1,996	21	1.04%				
28	B4galt4	6,818	6,787	31	0.45%				
28	Sppl3	9,786	9,748	38	0.39%				
28	Bax1	5,465	5,420	45	0.82%				
28	B4galt1	2,710	2,701	9	0.33%				
28	B4galt3	5,842	5,816	26	0.45%				
28	B4galt2	6,586	6,566	20	0.30%				
28	b4galt5	5,537	5,515	22	0.40%				
28	Bak1	10,467	10,387	80	0.76%				
29	GS	12,827	12,781	46	0.36%				
29	Tsta3	2,838	2,818	20	0.70%				
29	B4galt4	7,143	7,108	35	0.49%				
29	Sppl3	1,441	1,436	5	0.35%				
29	Bax1	6,520	6,464	56	0.86%				
29	B4galt1	215	213	2	0.93%				
29	B4galt3	5,944	5,926	18	0.30%				
29	B4galt2	7,788	7,774	14	0.18%				
29	b4galt5	6,839	6,814	25	0.37%				
29	Bak1	11,905	11,824	81	0.68%				

30	GS	13,189	13,147	42	0.32%				
30	Tsta3	2,225	2,184	41	1.84%				
30	B4galt4	7,758	7,664	94	1.21%				
30	Sppl3	11,607	11,552	55	0.47%				
30	Bax1	6,438	5,663	775	12.04%	-94			
30	B4galt1	1,395	1,389	6	0.43%				
30	B4galt3	5,847	5,799	48	0.82%				
30	B4galt2	7,905	7,884	21	0.27%				
30	b4galt5	6,962	6,914	48	0.69%				
30	Bak1	12,253	11,923	330	2.69%				
31	GS	15,818	15,767	51	0.32%				
31	Tsta3	3,228	3,217	11	0.34%				
31	B4galt4	9,048	8,982	66	0.73%				
31	Sppl3	14,024	13,934	90	0.64%				
31	Bax1	6,721	6,668	53	0.79%				
31	B4galt1	2,992	2,985	7	0.23%				
31	B4galt3	7,324	7,275	49	0.67%				
31	B4galt2	9,147	9,122	25	0.27%				
31	b4galt5	7,031	7,000	31	0.44%				
31	Bak1	15,452	15,343	109	0.71%				
32	GS	14,746	14,694	52	0.35%				
32	Tsta3	2,358	2,344	14	0.59%				
32	B4galt4	9,135	9,061	74	0.81%				
32	Sppl3	12,360	12,280	80	0.65%				
32	Bax1	5,763	5,705	58	1.01%				
32	B4galt1	2,750	2,735	15	0.55%				
32	B4galt3	7,724	7,688	36	0.47%				
32	B4galt2	9,385	9,366	19	0.20%				
32	b4galt5	6,342	6,312	30	0.47%				
32	Bak1	14,047	13,936	111	0.79%				
33	GS	16,703	16,641	62	0.37%				
33	Tsta3	1,957	1,943	14	0.72%				
33	B4galt4	11,233	11,180	53	0.47%				
33	Sppl3	15,936	15,867	69	0.43%				
33	Bax1	8,292	8,237	55	0.66%				
33	B4galt1	5,592	5,572	20	0.36%				
33	B4galt3	9,024	8,976	48	0.53%				
33	B4galt2	10,568	10,542	26	0.25%				
33	b4galt5	8,434	8,402	32	0.38%				
33	Bak1	15,366	15,245	121	0.79%				
34	GS	13,322	13,274	48	0.36%				
34	Tsta3	1,927	1,917	10	0.52%				
34	B4galt4	8,119	8,083	36	0.44%				
34	Sppl3	13,927	13,844	83	0.60%				
34	Bax1	5,687	5,639	48	0.84%				
34	B4galt1	4,136	4,116	20	0.48%				
34	B4galt3	6,909	6,868	41	0.59%				
34	B4galt2	7,685	7,662	23	0.30%				
34	b4galt5	6,273	6,246	27	0.43%				
34	Bak1	13,014	12,916	98	0.75%				

35	GS	12,819	12,764	55	0.43%				
35	Tsta3	1,224	1,217	7	0.57%				
35	B4galt4	7,950	7,903	47	0.59%				
35	Sppl3	11,439	11,375	64	0.56%				
35	Bax1	6,502	6,460	42	0.65%				
35	B4galt1	4,186	4,159	27	0.65%				
35	B4galt3	6,110	6,089	21	0.34%				
35	B4galt2	7,889	7,875	14	0.18%				
35	b4galt5	6,425	6,402	23	0.36%				
35	Bak1	11,675	11,583	92	0.79%				
36	GS	13,754	13,707	47	0.34%				
36	Tsta3	1,101	4	1,097	99.64%	-9 (inframe)	-2	1	
36	B4galt4	8,466	9	8,457	99.89%	-1	1		
36	Sppl3	15,272	15,174	98	0.64%				
36	Bax1	6,456	39	6,417	99.40%	-13	1		
36	B4galt1	5,012	4,996	16	0.32%				
36	B4galt3	7,643	5,721	1,922	25.15%	1			
36	B4galt2	8,321	4,212	4,109	49.38%	-9 (inframe)	1		
36	b4galt5	5,776	5,124	652	11.29%	-77			
36	Bak1	10,649	45	10,604	99.58%	-2	1		
37	GS	16,299	16,237	62	0.38%				
37	Tsta3	2,173	2,158	15	0.69%				
37	B4galt4	7,563	7,533	30	0.40%				
37	Sppl3	13,265	13,194	71	0.54%				
37	Bax1	4,713	4,669	44	0.93%				
37	B4galt1	5,965	5,939	26	0.44%				
37	B4galt3	6,112	6,082	30	0.49%				
37	B4galt2	8,505	8,480	25	0.29%				
37	b4galt5	8,436	8,402	34	0.40%				
37	Bak1	13,971	13,875	96	0.69%				
38	GS	13,767	13,712	55	0.40%				
38	Tsta3	1,956	1,945	11	0.56%				
38	B4galt4	6,257	6,232	25	0.40%				
38	Sppl3	9,435	9,389	46	0.49%				
38	Bax1	4,508	4,476	32	0.71%				
38	B4galt1	3,699	3,692	7	0.19%				
38	B4galt3	4,928	4,907	21	0.43%				
38	B4galt2	6,308	6,289	19	0.30%				
38	b4galt5	7,178	7,145	33	0.46%				
38	Bak1	11,411	11,343	68	0.60%				
39	GS	14,891	14,850	41	0.28%				
39	Tsta3	2,001	1,992	9	0.45%				
39	B4galt4	7,236	7,213	23	0.32%				
39	Sppl3	3,546	3,533	13	0.37%				
39	Bax1	7,412	7,125	287	3.87%				
39	B4galt1	358	357	1	0.28%				
39	B4galt3	5,049	5,024	25	0.50%				
39	B4galt2	7,909	5,746	2,163	27.35%	-7			
39	b4galt5	7,493	7,457	36	0.48%				
39	Bak1	13,375	13,289	86	0.64%				
40	GS	16,644	16,590	54	0.32%				
40	Tsta3	1,758	3	1,755	99.83%	-16	-6 (inframe)		
40	B4galt4	7,420	12	7,408	99.84%	-4	-1	1	
40	Sppl3	11,969	3,521	8,448	70.58%	-27 (inframe)	-17	-12 (inframe)	
40	Bax1	7,209	3,904	3,305	45.85%	1			
40	B4galt1	7,925	7,897	28	0.35%				
40	B4galt3	5,687	15	5,672	99.74%	-38	-9 (inframe)	1	
40	B4galt2	7,788	3,390	4,398	56.47%	-25	-18 (inframe)	-14	
40	b4galt5	6,956	5,540	1,416	20.36%	-3 (inframe)			
40	Bak1	12,906	28	12,878	99.78%	-6 (inframe)	-3 (inframe)	-2	

41	GS	15,903	15,843	60	0.38%				
41	Tsta3	2,195	2,174	21	0.96%				
41	B4galt4	7,309	7,271	38	0.52%				
41	Sppl3	10,015	9,952	63	0.63%				
41	Bax1	7,835	7,788	47	0.60%				
41	B4galt1	8,538	8,493	45	0.53%				
41	B4galt3	5,898	5,867	31	0.53%				
41	B4galt2	7,526	7,505	21	0.28%				
41	b4galt5	8,003	7,955	48	0.60%				
41	Bak1	14,015	13,935	80	0.57%				
42	GS	13,075	13,026	49	0.37%				
42	Tsta3	1,865	1,856	9	0.48%				
42	B4galt4	6,308	6,276	32	0.51%				
42	Sppl3	8,665	8,629	36	0.42%				
42	Bax1	5,855	5,807	48	0.82%				
42	B4galt1	1,773	1,769	4	0.23%				
42	B4galt3	5,136	5,112	24	0.47%				
42	B4galt2	5,947	5,927	20	0.34%				
42	b4galt5	5,756	5,731	25	0.43%				
42	Bak1	10,658	10,581	77	0.72%				
43	GS	12,773	12,724	49	0.38%				
43	Tsta3	1,484	1,479	5	0.34%				
43	B4galt4	6,550	6,529	21	0.32%				
43	Sppl3	10,584	10,520	64	0.60%				
43	Bax1	6,005	5,958	47	0.78%				
43	B4galt1	2,875	2,870	5	0.17%				
43	B4galt3	4,709	4,690	19	0.40%				
43	B4galt2	6,649	6,634	15	0.23%				
43	b4galt5	6,385	6,360	25	0.39%				
43	Bak1	11,125	11,054	71	0.64%				
44	GS	12,924	12,880	44	0.34%				
44	Tsta3	1,416	1,405	11	0.78%				
44	B4galt4	7,033	6,985	48	0.68%				
44	Sppl3	11,268	11,200	68	0.60%				
44	Bax1	6,686	6,638	48	0.72%				
44	B4galt3	5,019	4,997	22	0.44%				
44	B4galt2	6,660	6,642	18	0.27%				
44	b4galt5	6,789	6,757	32	0.47%				
44	Bak1	10,699	10,623	76	0.71%				
45	GS	17,172	17,108	64	0.37%				
45	Tsta3	2,052	2,042	10	0.49%				
45	B4galt4	7,919	7,887	32	0.40%				
45	Sppl3	13,427	13,368	59	0.44%				
45	Bax1	7,130	7,082	48	0.67%				
45	B4galt1	8,365	8,332	33	0.39%				
45	B4galt3	5,931	5,901	30	0.51%				
45	B4galt2	7,095	7,071	24	0.34%				
45	b4galt5	7,403	7,365	38	0.51%				
45	Bak1	13,909	13,826	83	0.60%				

46	GS	13,631	13,586	45	0.33%				
46	Tsta3	2,068	2,063	5	0.24%				
46	B4galt4	6,610	6,573	37	0.56%				
46	Sppl3	10,990	10,942	48	0.44%				
46	Bax1	6,046	5,992	54	0.89%				
46	B4galt1	7,131	7,106	25	0.35%				
46	B4galt3	5,638	5,610	28	0.50%				
46	B4galt2	6,424	6,411	13	0.20%				
46	b4galt5	5,455	5,430	25	0.46%				
46	Bak1	11,590	11,509	81	0.70%				
47	GS	8,422	8,396	26	0.31%				
47	Tsta3	972	966	6	0.62%				
47	B4galt4	3,462	3,442	20	0.58%				
47	Sppl3	9,833	9,772	61	0.62%				
47	Bax1	5,667	5,619	48	0.85%				
47	B4galt1	4,584	4,570	14	0.31%				
47	B4galt3	3,684	3,672	12	0.33%				
47	B4galt2	5,385	5,367	18	0.33%				
47	b4galt5	3,373	3,356	17	0.50%				
47	Bak1	8,304	8,250	54	0.65%				
48	GS	17,381	17,321	60	0.35%				
48	Tsta3	375	372	3	0.80%				
48	B4galt4	8,258	8,219	39	0.47%				
48	Sppl3	16,435	16,348	87	0.53%				
48	Bax1	8,990	8,921	69	0.77%				
48	B4galt3	6,123	6,097	26	0.42%				
48	B4galt2	7,039	7,025	14	0.20%				
48	b4galt5	9,135	9,083	52	0.57%				
48	Bak1	14,016	13,911	105	0.75%				
49	GS	16,450	16,397	53	0.32%				
49	Tsta3	762	5	757	99.34%	-16	-2	-1	
49	B4galt4	7,009	3,485	3,524	50.28%	-1			
49	Sppl3	11,926	11,838	88	0.74%				
49	Bax1	6,849	19	6,830	99.72%	-26	-9 (inframe)		
49	B4galt1	463	460	3	0.65%				
49	B4galt3	4,972	2,577	2,395	48.17%	-2			
49	B4galt2	6,251	6,240	11	0.18%				
49	b4galt5	6,300	5,450	850	13.49%	1			
49	Bak1	6,497	48	6,449	99.26%	-14			
50	GS	11,209	11,177	32	0.29%				
50	Tsta3	656	650	6	0.91%				
50	B4galt4	4,887	4,870	17	0.35%				
50	Sppl3	7,780	7,746	34	0.44%				
50	Bax1	3,272	3,250	22	0.67%				
50	B4galt1	4,944	4,923	21	0.42%				
50	B4galt3	4,006	3,988	18	0.45%				
50	B4galt2	4,121	4,108	13	0.32%				
50	b4galt5	4,526	4,506	20	0.44%				
50	Bak1	8,400	8,347	53	0.63%				
51	GS	14,479	14,429	50	0.35%				
51	Tsta3	1,494	1,485	9	0.60%				
51	B4galt4	6,163	6,133	30	0.49%				
51	Sppl3	5,785	5,759	26	0.45%				
51	Bax1	5,391	2,465	2,926	54.28%	-15 (inframe)	-3 (inframe)		
51	B4galt1	5,306	5,288	18	0.34%				
51	B4galt3	4,406	4,385	21	0.48%				
51	B4galt2	5,652	5,635	17	0.30%				
51	b4galt5	5,797	5,773	24	0.41%				
51	Bak1	11,691	11,415	276	2.36%				

52	GS	11,756	11,718	38	0.32%				
52	Tsta3	1,038	1,031	7	0.67%				
52	B4galt4	5,327	5,298	29	0.54%				
52	Sppl3	3,267	3,252	15	0.46%				
52	Bax1	3,915	3,881	34	0.87%				
52	B4galt1	5,034	5,024	10	0.20%				
52	B4galt3	4,196	4,176	20	0.48%				
52	B4galt2	4,388	4,377	11	0.25%				
52	b4galt5	4,255	4,234	21	0.49%				
52	Bak1	8,711	8,658	53	0.61%				
53	GS	10,832	10,788	44	0.41%				
53	Tsta3	996	994	2	0.20%				
53	B4galt4	4,791	4,755	36	0.75%				
53	Sppl3	6,091	6,067	24	0.39%				
53	Bax1	4,276	4,234	42	0.98%				
53	B4galt1	4,792	4,771	21	0.44%				
53	B4galt3	3,439	3,411	28	0.81%				
53	B4galt2	4,208	4,202	6	0.14%				
53	b4galt5	4,071	4,050	21	0.52%				
53	Bak1	8,542	8,479	63	0.74%				
54	GS	13,659	13,626	33	0.24%				
54	Tsta3	1,321	1,309	12	0.91%				
54	B4galt4	6,468	6,436	32	0.49%				
54	Sppl3	6,244	6,217	27	0.43%				
54	Bax1	5,917	5,853	64	1.08%				
54	B4galt1	3,167	3,154	13	0.41%				
54	B4galt3	5,430	5,402	28	0.52%				
54	B4galt2	5,194	5,185	9	0.17%				
54	b4galt5	4,799	4,773	26	0.54%				
54	Bak1	10,556	10,485	71	0.67%				
55	GS	10,770	10,731	39	0.36%				
55	Tsta3	934	927	7	0.75%				
55	B4galt4	4,448	4,434	14	0.31%				
55	Sppl3	6,443	6,418	25	0.39%				
55	Bax1	4,241	4,210	31	0.73%				
55	B4galt1	4,007	3,992	15	0.37%				
55	B4galt3	3,402	3,389	13	0.38%				
55	B4galt2	3,677	3,660	17	0.46%				
55	b4galt5	3,565	3,548	17	0.48%				
55	Bak1	8,960	8,896	64	0.71%				
56	GS	12,172	12,130	42	0.35%				
56	Tsta3	739	280	459	62.11%	-9 (inframe)	-6 (inframe)	-2	-1
56	B4galt4	5,513	586	4,927	89.37%	-2	-1	1	
56	Sppl3	8,514	8,461	53	0.62%				
56	Bax1	4,789	3,405	1,384	28.90%	-1			
56	B4galt3	4,358	3,627	731	16.77%	-10			
56	B4galt2	4,323	3,708	615	14.23%	44			
56	b4galt5	4,800	4,782	18	0.38%				
56	Bak1	9,713	9,641	72	0.74%				

57	GS	10,381	10,339	42	0.40%				
57	Tsta3	1,028	3	1,025	99.71%	-11	1		
57	B4galt4	4,956	2,433	2,523	50.91%	-1			
57	Sppl3	8,063	8,025	38	0.47%				
57	Bax1	4,577	16	4,561	99.65%	-1			
57	B4galt1	2,436	2,425	11	0.45%				
57	B4galt3	3,831	3,807	24	0.63%				
57	B4galt2	4,431	3,882	549	12.39%	-18 (inframe)			
57	b4galt5	4,356	4,334	22	0.51%				
57	Bak1	8,750	8,654	96	1.10%				
58	GS	11,678	11,618	60	0.51%				
58	Tsta3	684	678	6	0.88%				
58	B4galt4	5,164	5,140	24	0.46%				
58	Sppl3	9,177	9,127	50	0.54%				
58	Bax1	3,786	3,757	29	0.77%				
58	B4galt1	4,742	4,730	12	0.25%				
58	B4galt3	3,564	3,548	16	0.45%				
58	B4galt2	3,833	3,823	10	0.26%				
58	b4galt5	4,343	4,327	16	0.37%				
58	Bak1	9,251	9,186	65	0.70%				
59	GS	11,225	11,185	40	0.36%				
59	Tsta3	730	726	4	0.55%				
59	B4galt4	5,241	5,227	14	0.27%				
59	Sppl3	8,751	8,701	50	0.57%				
59	Bax1	4,472	4,430	42	0.94%				
59	B4galt1	4,850	4,836	14	0.29%				
59	B4galt3	3,902	3,873	29	0.74%				
59	B4galt2	4,097	4,088	9	0.22%				
59	b4galt5	4,176	4,153	23	0.55%				
59	Bak1	8,780	8,708	72	0.82%				
60	GS	13,981	13,921	60	0.43%				
60	Tsta3	147	147	0	0.00%				
60	B4galt4	6,155	6,123	32	0.52%				
60	Sppl3	12,536	12,454	82	0.65%				
60	Bax1	7,368	55	7,313	99.25%	1			
60	B4galt3	4,027	4,010	17	0.42%				
60	B4galt2	5,598	5,583	15	0.27%				
60	b4galt5	5,816	5,792	24	0.41%				
60	Bak1	10,753	6,474	4,279	39.79%	-64			
61	GS	13,448	13,400	48	0.36%				
61	Tsta3	3,109	3,091	18	0.58%				
61	B4galt4	7,330	7,301	29	0.40%				
61	Sppl3	12,239	12,166	73	0.60%				
61	Bax1	6,378	6,335	43	0.67%				
61	B4galt1	6,559	6,538	21	0.32%				
61	B4galt3	3,468	3,452	16	0.46%				
61	B4galt2	8,389	8,362	27	0.32%				
61	b4galt5	7,138	7,097	41	0.57%				
61	Bak1	11,887	11,798	89	0.75%				
62	GS	11,394	11,360	34	0.30%				
62	Tsta3	2,228	2,217	11	0.49%				
62	B4galt4	5,849	5,827	22	0.38%				
62	Sppl3	9,374	9,322	52	0.55%				
62	Bax1	4,767	4,727	40	0.84%				
62	B4galt1	5,131	5,113	18	0.35%				
62	B4galt3	3,908	3,901	7	0.18%				
62	B4galt2	5,627	5,614	13	0.23%				
62	b4galt5	4,894	4,878	16	0.33%				
62	Bak1	9,493	9,436	57	0.60%				

63	GS	12,980	12,943	37	0.29%				
63	Tsta3	1,919	1,907	12	0.63%				
63	B4galt4	6,551	6,512	39	0.60%				
63	Sppl3	7,894	7,856	38	0.48%				
63	Bax1	6,597	6,547	50	0.76%				
63	B4galt1	5,429	5,409	20	0.37%				
63	B4galt3	4,771	4,754	17	0.36%				
63	B4galt2	7,520	7,503	17	0.23%				
63	b4galt5	5,892	5,871	21	0.36%				
63	Bak1	11,382	11,315	67	0.59%				
64	GS	12,356	12,309	47	0.38%				
64	Tsta3	1,096	1,093	3	0.27%				
64	B4galt4	6,534	6,504	30	0.46%				
64	Bax1	6,209	6,151	58	0.93%				
64	B4galt1	6,803	6,778	25	0.37%				
64	B4galt3	5,375	5,349	26	0.48%				
64	B4galt2	6,767	6,740	27	0.40%				
64	b4galt5	4,748	4,730	18	0.38%				
64	Bak1	10,529	10,459	70	0.66%				
65	GS	11,179	11,137	42	0.38%				
65	Tsta3	1,011	1,006	5	0.49%				
65	B4galt4	6,006	5,960	46	0.77%				
65	Sppl3	7,830	7,797	33	0.42%				
65	Bax1	5,933	5,902	31	0.52%				
65	B4galt1	5,245	5,219	26	0.50%				
65	B4galt3	4,998	4,962	36	0.72%				
65	B4galt2	6,378	6,358	20	0.31%				
65	b4galt5	4,752	4,724	28	0.59%				
65	Bak1	9,987	9,925	62	0.62%				
66	GS	12,825	12,776	49	0.38%				
66	Tsta3	867	863	4	0.46%				
66	B4galt4	6,164	6,143	21	0.34%				
66	Sppl3	6,043	6,016	27	0.45%				
66	Bax1	5,715	5,665	50	0.87%				
66	B4galt1	5,939	5,915	24	0.40%				
66	B4galt3	4,930	4,904	26	0.53%				
66	B4galt2	6,548	6,534	14	0.21%				
66	b4galt5	5,021	4,992	29	0.58%				
66	Bak1	10,273	10,205	68	0.66%				
67	GS	11,375	11,329	46	0.40%				
67	Tsta3	1,150	1,145	5	0.43%				
67	B4galt4	5,422	5,402	20	0.37%				
67	Sppl3	8,671	8,618	53	0.61%				
67	Bax1	5,771	5,728	43	0.75%				
67	B4galt1	5,047	5,031	16	0.32%				
67	B4galt3	3,914	3,897	17	0.43%				
67	B4galt2	6,165	6,151	14	0.23%				
67	b4galt5	4,885	4,856	29	0.59%				
67	Bak1	10,167	10,098	69	0.68%				
68	GS	8,941	8,910	31	0.35%				

68	Tsta3	261	1	260	99.62%	-6 (inframe)			
68	B4galt4	3,268	17	3,251	99.48%	-9 (inframe)	1		
68	Sppl3	9,359	53	9,306	99.43%	-42 (inframe)	1		
68	Bax1	2,484	23	2,461	99.07%	1			
68	B4galt1	3,152	3,143	9	0.29%				
68	B4galt3	3,184	13	3,171	99.59%	-10	1		
68	B4galt2	3,753	10	3,743	99.73%	-21 (inframe)	-2		
68	b4galt5	2,478	3	2,475	99.88%	-4	-3 (inframe)		
68	Bak1	7,745	31	7,714	99.60%	1	2		
69	GS	13,329	13,282	47	0.35%				
69	Tsta3	450	449	1	0.22%				
69	B4galt4	6,049	6,019	30	0.50%				
69	Sppl3	5,129	5,109	20	0.39%				
69	Bax1	6,459	6,418	41	0.63%				
69	B4galt1	6,061	6,036	25	0.41%				
69	B4galt3	4,994	4,974	20	0.40%				
69	B4galt2	6,530	6,512	18	0.28%				
69	b4galt5	5,547	5,525	22	0.40%				
69	Bak1	11,772	11,703	69	0.59%				
70	GS	15,731	15,673	58	0.37%				
70	Tsta3	1,314	1,310	4	0.30%				
70	B4galt4	7,633	7,602	31	0.41%				
70	Sppl3	12,137	12,064	73	0.60%				
70	Bax1	7,008	6,962	46	0.66%				
70	B4galt1	7,539	7,509	30	0.40%				
70	B4galt3	5,974	5,942	32	0.54%				
70	B4galt2	7,659	7,640	19	0.25%				
70	b4galt5	6,299	6,270	29	0.46%				
70	Bak1	13,123	13,024	99	0.75%				
71	GS	11,945	11,897	48	0.40%				
71	Tsta3	718	714	4	0.56%				
71	B4galt4	6,061	6,039	22	0.36%				
71	Sppl3	9,218	9,164	54	0.59%				
71	Bax1	5,524	5,473	51	0.92%				
71	B4galt1	5,506	5,481	25	0.45%				
71	B4galt3	5,169	5,138	31	0.60%				
71	B4galt2	6,691	6,670	21	0.31%				
71	b4galt5	4,687	4,669	18	0.38%				
71	Bak1	10,580	10,517	63	0.60%				
72	GS	16,144	16,088	56	0.35%				
72	Tsta3	1,026	540	486	47.37%	-6 (inframe)	-1	1	
72	B4galt4	6,164	4,248	1,916	31.08%	-8	2		
72	Sppl3	14,322	13,920	402	2.81%				
72	Bax1	5,611	3,905	1,706	30.40%	-2			
72	B4galt1	4,362	4,348	14	0.32%				
72	B4galt3	5,923	5,893	30	0.51%				
72	B4galt2	5,987	4,214	1,773	29.61%	-18 (inframe)	-10	36 (inframe)	
72	b4galt5	4,523	4,500	23	0.51%				
72	Bak1	12,621	8,072	4,549	36.04%	-24 (inframe)	-2		
73	GS	14,383	14,327	56	0.39%				
73	Tsta3	155	153	2	1.29%				
73	B4galt4	7,040	15	7,025	99.79%	-1	1		
73	Sppl3	13,827	13,752	75	0.54%				
73	Bax1	5,828	5,028	800	13.73%	-1			
73	B4galt1	6,462	6,439	23	0.36%				
73	B4galt3	5,039	5,014	25	0.50%				
73	B4galt2	7,892	7,873	19	0.24%				
73	b4galt5	7,767	6,137	1,630	20.99%	-3 (inframe)			
73	Bak1	11,197	8,966	2,231	19.92%	-10			

74	GS	13,001	12,955	46	0.35%				
74	Tsta3	540	538	2	0.37%				
74	B4galt4	5,712	5,688	24	0.42%				
74	Sppl3	7,049	7,026	23	0.33%				
74	Bax1	4,796	4,758	38	0.79%				
74	B4galt1	6,551	6,530	21	0.32%				
74	B4galt3	5,541	5,509	32	0.58%				
74	B4galt2	6,827	6,807	20	0.29%				
74	b4galt5	5,216	5,195	21	0.40%				
74	Bak1	11,151	11,085	66	0.59%				
75	GS	13,659	13,602	57	0.42%				
75	Tsta3	1,468	1,457	11	0.75%				
75	B4galt4	6,132	6,113	19	0.31%				
75	Sppl3	12,046	11,978	68	0.56%				
75	Bax1	6,073	6,025	48	0.79%				
75	B4galt1	6,853	6,831	22	0.32%				
75	B4galt3	5,388	5,358	30	0.56%				
75	B4galt2	6,697	6,681	16	0.24%				
75	b4galt5	6,158	6,133	25	0.41%				
75	Bak1	11,851	11,776	75	0.63%				
76	GS	13,240	13,189	51	0.39%				
76	Tsta3	1,471	1,467	4	0.27%				
76	B4galt4	5,484	8	5,476	99.85%	-11			
76	Sppl3	7,559	7,534	25	0.33%				
76	Bax1	5,498	49	5,449	99.11%	1			
76	B4galt1	6,986	6,949	37	0.53%				
76	B4galt3	5,349	5,322	27	0.50%				
76	B4galt2	6,207	6,193	14	0.23%				
76	b4galt5	6,312	4,985	1,327	21.02%	-15 (inframe)			
76	Bak1	11,298	9,135	2,163	19.14%	17			
77	GS	15,658	15,610	48	0.31%				
77	Tsta3	1,611	489	1,122	69.65%	-12 (inframe)	-6 (inframe)		
77	B4galt4	6,840	6,814	26	0.38%				
77	Sppl3	8,975	8,922	53	0.59%				
77	Bax1	6,734	10	6,724	99.85%	2			
77	B4galt1	7,754	7,724	30	0.39%				
77	B4galt3	6,069	6,047	22	0.36%				
77	B4galt2	6,571	6,556	15	0.23%				
77	b4galt5	6,623	6,588	35	0.53%				
77	Bak1	12,019	61	11,958	99.49%	-8	1		
78	GS	13,972	13,919	53	0.38%				
78	Tsta3	1,540	3	1,537	99.81%	-16	-9 (inframe)	-6 (inframe)	-2
78	B4galt4	6,129	11	6,118	99.82%	-1			
78	Sppl3	6,470	6,437	33	0.51%				
78	Bax1	6,508	3,289	3,219	49.46%	-1			
78	B4galt1	7,163	7,137	26	0.36%				
78	B4galt3	6,074	3,762	2,312	38.06%	-7			
78	B4galt2	6,903	2,002	4,901	71.00%	-19	-8		
78	b4galt5	6,596	3,878	2,718	41.21%	1			
78	Bak1	10,695	46	10,649	99.57%	-13	-2	1	

79	GS	13,040	12,999	41	0.31%				
79	Tsta3	916	912	4	0.44%				
79	B4galt4	6,398	6,370	28	0.44%				
79	Sppl3	5,473	5,447	26	0.48%				
79	Bax1	6,127	6,079	48	0.78%				
79	B4galt1	6,448	6,427	21	0.33%				
79	B4galt3	5,313	5,291	22	0.41%				
79	B4galt2	6,924	6,907	17	0.25%				
79	b4galt5	5,379	5,353	26	0.48%				
79	Bak1	11,046	10,973	73	0.66%				
80	GS	17,245	17,196	49	0.28%				
80	Tsta3	1,957	1,946	11	0.56%				
80	B4galt4	8,131	8,101	30	0.37%				
80	Sppl3	12,279	12,229	50	0.41%				
80	Bax1	5,989	5,948	41	0.68%				
80	B4galt1	8,079	8,042	37	0.46%				
80	B4galt3	6,361	6,336	25	0.39%				
80	B4galt2	8,249	8,228	21	0.25%				
80	b4galt5	7,183	7,141	42	0.58%				
80	Bak1	13,892	13,777	115	0.83%				
81	GS	14,715	14,661	54	0.37%				
81	Tsta3	1,414	1,405	9	0.64%				
81	B4galt4	7,478	7,446	32	0.43%				
81	Sppl3	5,825	5,792	33	0.57%				
81	Bax1	6,327	6,269	58	0.92%				
81	B4galt1	7,377	7,350	27	0.37%				
81	B4galt3	5,942	5,907	35	0.59%				
81	B4galt2	7,139	7,123	16	0.22%				
81	b4galt5	6,511	6,480	31	0.48%				
81	Bak1	12,351	12,255	96	0.78%				
82	GS	15,670	15,610	60	0.38%				
82	Tsta3	1,738	1,732	6	0.35%				
82	B4galt4	7,787	7,758	29	0.37%				
82	Sppl3	11,730	11,672	58	0.49%				
82	Bax1	6,278	6,233	45	0.72%				
82	B4galt1	7,261	7,234	27	0.37%				
82	B4galt3	6,439	6,408	31	0.48%				
82	B4galt2	7,372	7,348	24	0.33%				
82	b4galt5	6,977	6,949	28	0.40%				
82	Bak1	13,021	12,917	104	0.80%				
83	GS	12,405	12,351	54	0.44%				
83	Tsta3	865	858	7	0.81%				
83	B4galt4	5,757	5,733	24	0.42%				
83	Sppl3	11,352	11,188	164	1.44%				
83	Bax1	4,700	4,661	39	0.83%				
83	B4galt1	5,196	5,171	25	0.48%				
83	B4galt3	4,814	4,800	14	0.29%				
83	B4galt2	5,088	5,077	11	0.22%				
83	b4galt5	3,906	3,893	13	0.33%				
83	Bak1	10,701	10,619	82	0.77%				
84	GS	17,766	17,683	83	0.47%				
84	Tsta3	1,140	1,135	5	0.44%				
84	B4galt4	9,780	9,748	32	0.33%				
84	Sppl3	16,357	16,243	114	0.70%				
84	Bax1	9,036	4,673	4,363	48.28%	1			
84	B4galt1	9,155	9,122	33	0.36%				
84	B4galt3	7,873	7,832	41	0.52%				
84	B4galt2	9,303	9,268	35	0.38%				
84	b4galt5	7,959	7,922	37	0.46%				
84	Bak1	14,959	14,852	107	0.72%				

85	GS	13,950	13,904	46	0.33%				
85	Tsta3	2,290	2,273	17	0.74%				
85	B4galt4	6,277	6,251	26	0.41%				
85	Sppl3	12,210	12,145	65	0.53%				
85	Bax1	7,641	7,587	54	0.71%				
85	B4galt1	7,475	7,448	27	0.36%				
85	B4galt3	4,857	4,838	19	0.39%				
85	B4galt2	6,806	6,783	23	0.34%				
85	b4galt5	8,982	8,938	44	0.49%				
85	Bak1	12,152	12,068	84	0.69%				
86	GS	6,450	6,424	26	0.40%				
86	Tsta3	888	884	4	0.45%				
86	B4galt4	2,075	2,066	9	0.43%				
86	Sppl3	10,017	9,799	218	2.18%				
86	Bax1	4,338	4,304	34	0.78%				
86	B4galt1	3,930	3,915	15	0.38%				
86	B4galt3	2,760	2,749	11	0.40%				
86	B4galt2	4,335	4,326	9	0.21%				
86	b4galt5	3,419	3,403	16	0.47%				
86	Bak1	6,770	6,717	53	0.78%				
87	GS	15,132	15,085	47	0.31%				
87	Tsta3	1,109	1,105	4	0.36%				
87	B4galt4	8,043	8,009	34	0.42%				
87	Sppl3	13,565	13,498	67	0.49%				
87	Bax1	8,000	7,955	45	0.56%				
87	B4galt1	7,638	7,612	26	0.34%				
87	B4galt3	5,980	5,958	22	0.37%				
87	B4galt2	8,122	8,093	29	0.36%				
87	b4galt5	8,882	8,844	38	0.43%				
87	Bak1	13,282	13,194	88	0.66%				
88	GS	13,320	13,268	52	0.39%				
88	Tsta3	1,393	7	1,386	99.50%	-12 (inframe)	1		
88	B4galt4	6,261	7	6,254	99.89%	-1			
88	Sppl3	9,264	9,206	58	0.63%				
88	Bax1	6,787	6	6,781	99.91%	-2			
88	B4galt1	1,742	1,733	9	0.52%				
88	B4galt3	5,443	1,859	3,584	65.85%	1			
88	B4galt2	7,810	7,788	22	0.28%				
88	b4galt5	7,123	7,094	29	0.41%				
88	Bak1	12,230	36	12,194	99.71%	-2	1		
89	GS	15,272	15,226	46	0.30%				
89	Tsta3	1,626	1,448	178	10.95%	-16			
89	B4galt4	7,765	6,706	1,059	13.64%	1			
89	Sppl3	16,715	16,613	102	0.61%				
89	Bax1	4,798	4,755	43	0.90%				
89	B4galt1	8,391	8,352	39	0.46%				
89	B4galt3	6,132	6,106	26	0.42%				
89	B4galt2	7,488	7,471	17	0.23%				
89	b4galt5	9,534	9,497	37	0.39%				
89	Bak1	15,688	15,591	97	0.62%				

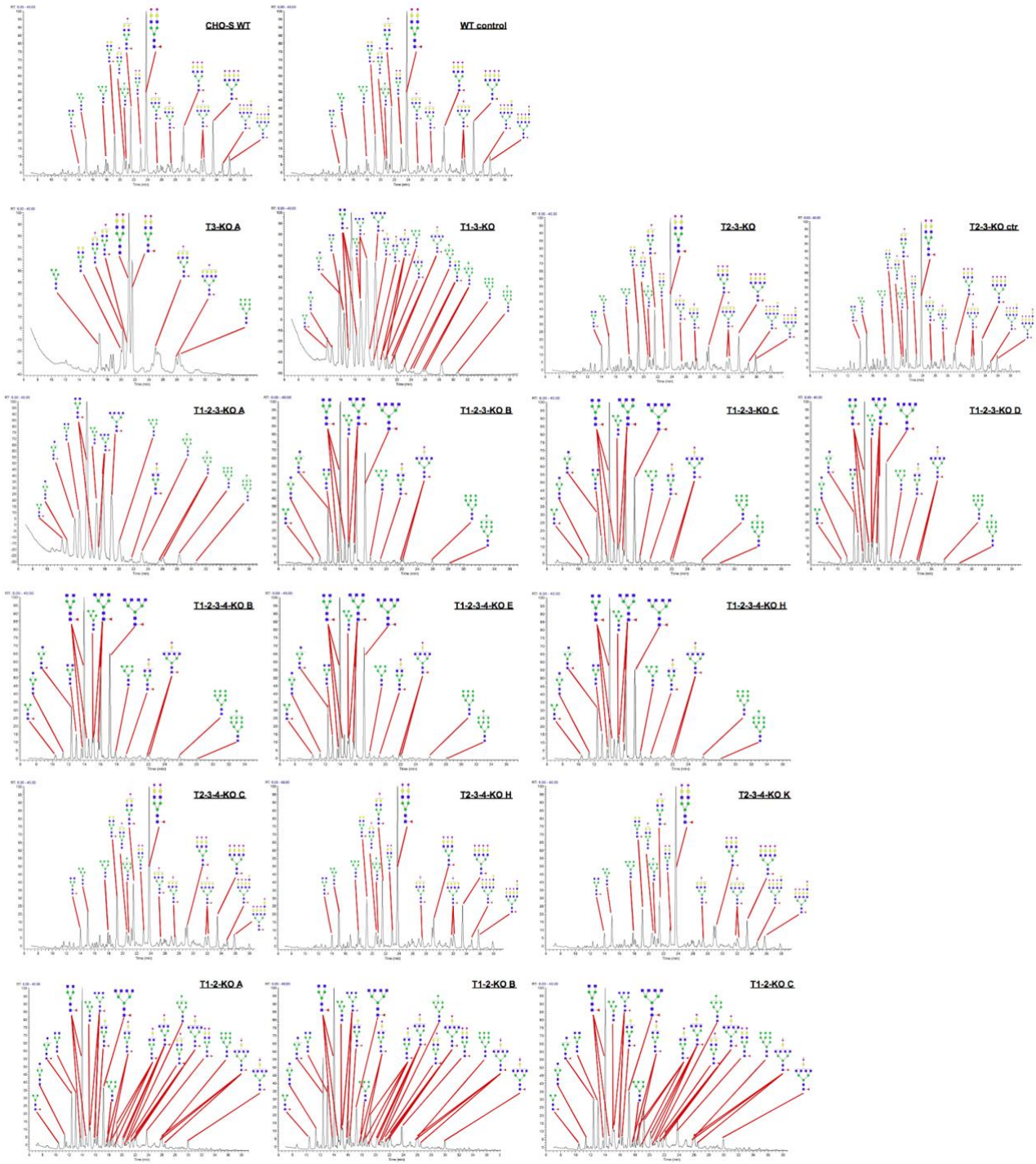
90	GS	14,779	14,728	51	0.35%				
90	Tsta3	1,838	0	1,838	100.00%	-6 (inframe)	-4		
90	B4galt4	4,615	25	4,590	99.46%	1	2		
90	Sppl3	15,433	8,365	7,068	45.80%	-69 (inframe)	-12 (inframe)	-8	1
90	Bax1	7,494	6	7,488	99.92%	-16	-2		
90	B4galt1	4,277	4,260	17	0.40%				
90	B4galt3	5,970	20	5,950	99.66%	1	2		
90	B4galt2	9,600	920	8,680	90.42%	-54 (inframe)	-18 (inframe)	-17	
90	b4galt5	9,428	3,993	5,435	57.65%	-16	-3 (inframe)	3 (inframe)	
90	Bak1	15,423	66	15,357	99.57%	-5	-3 (inframe)	1	
91	GS	18,594	18,525	69	0.37%				
91	Tsta3	1,367	1,356	11	0.80%				
91	B4galt4	8,177	8,141	36	0.44%				
91	Sppl3	10,692	10,603	89	0.83%				
91	Bax1	9,211	5	9,206	99.95%	-15 (inframe)			
91	B4galt1	7,542	7,508	34	0.45%				
91	B4galt3	6,997	5,364	1,633	23.34%	1			
91	B4galt2	8,854	8,838	16	0.18%				
91	Bak1	15,101	14,701	400	2.65%				
92	GS	16,293	16,230	63	0.39%				
92	Tsta3	1,342	995	347	25.86%	-2			
92	B4galt4	7,442	16	7,426	99.79%	-1	1		
92	Sppl3	15,551	15,427	124	0.80%				
92	Bax1	7,042	42	7,000	99.40%	-2	1		
92	B4galt1	5,390	5,361	29	0.54%				
92	B4galt3	6,459	6,434	25	0.39%				
92	B4galt2	7,731	7,708	23	0.30%				
92	Bak1	13,296	9,955	3,341	25.13%	-13			
93	GS	20,260	20,188	72	0.36%				
93	Tsta3	1,178	1,169	9	0.76%				
93	B4galt4	8,714	8,680	34	0.39%				
93	Sppl3	17,827	17,708	119	0.67%				
93	Bax1	9,377	13	9,364	99.86%	-1			
93	B4galt1	1,337	1,337	0	0.00%				
93	B4galt3	8,125	8,093	32	0.39%				
93	B4galt2	9,835	9,816	19	0.19%				
93	b4galt5	9,874	9,825	49	0.50%				
93	Bak1	16,977	13,559	3,418	20.13%	2			

Supplementary Material for

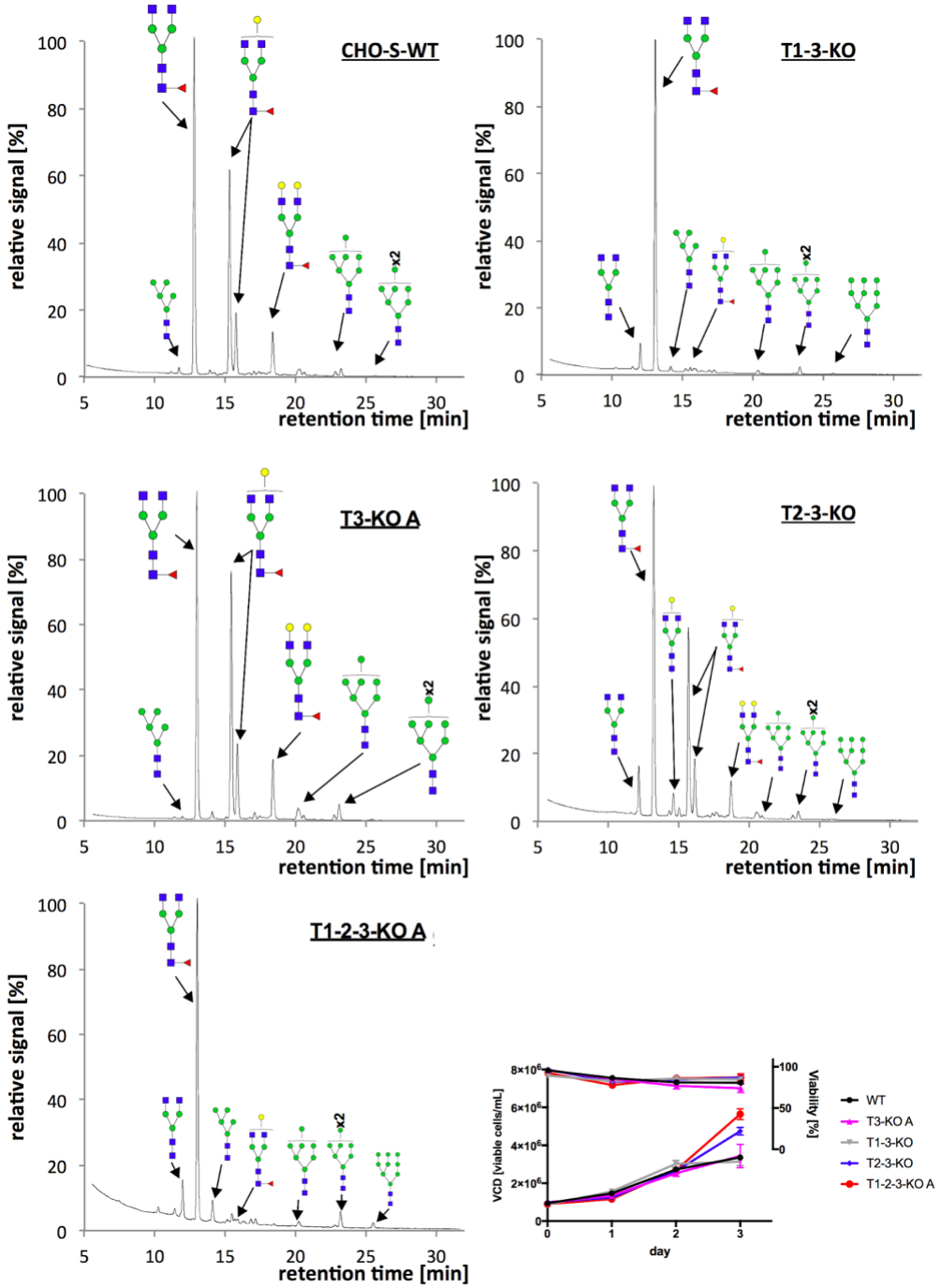
CRISPR/Cas9-multiplexed editing of Chinese hamster ovary B4Gal-T1, 2, 3 and 4 tailors N-glycan profiles of therapeutics and secreted host cell proteins

Thomas Amann[1], Anders Holmgaard Hansen[1], Stefan Kol[1], Gyun Min Lee[1,2], Mikael Rørdam Andersen[3], Helene Faustrup Kildegaard[1]

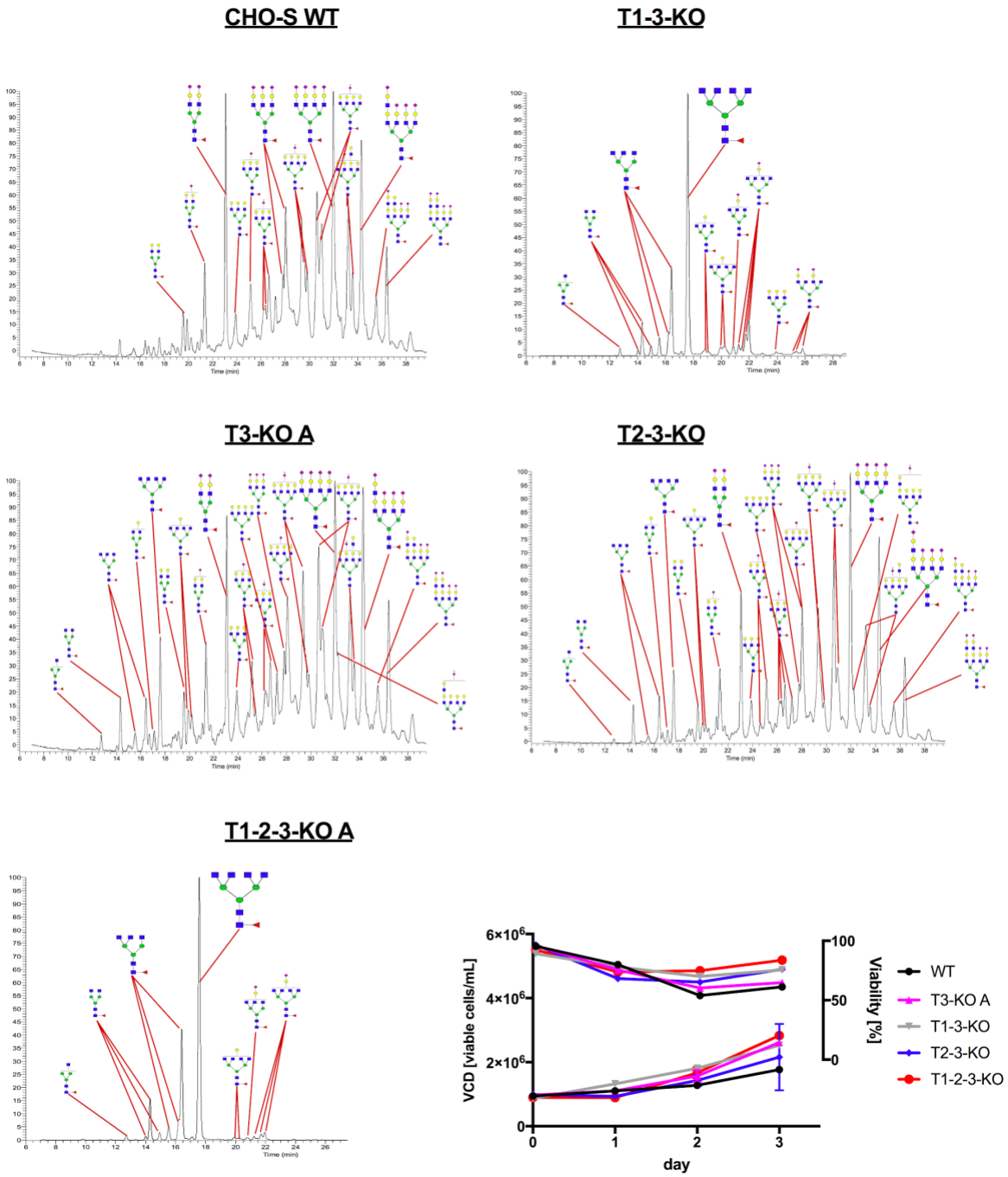
[1][The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Lyngby, Denmark]
[2][Department of Biological Sciences, KAIST, Daejeon, Republic of Korea]
[3][Department of Biotechnology and Biomedicine, Technical University of Denmark, Lyngby, Denmark]



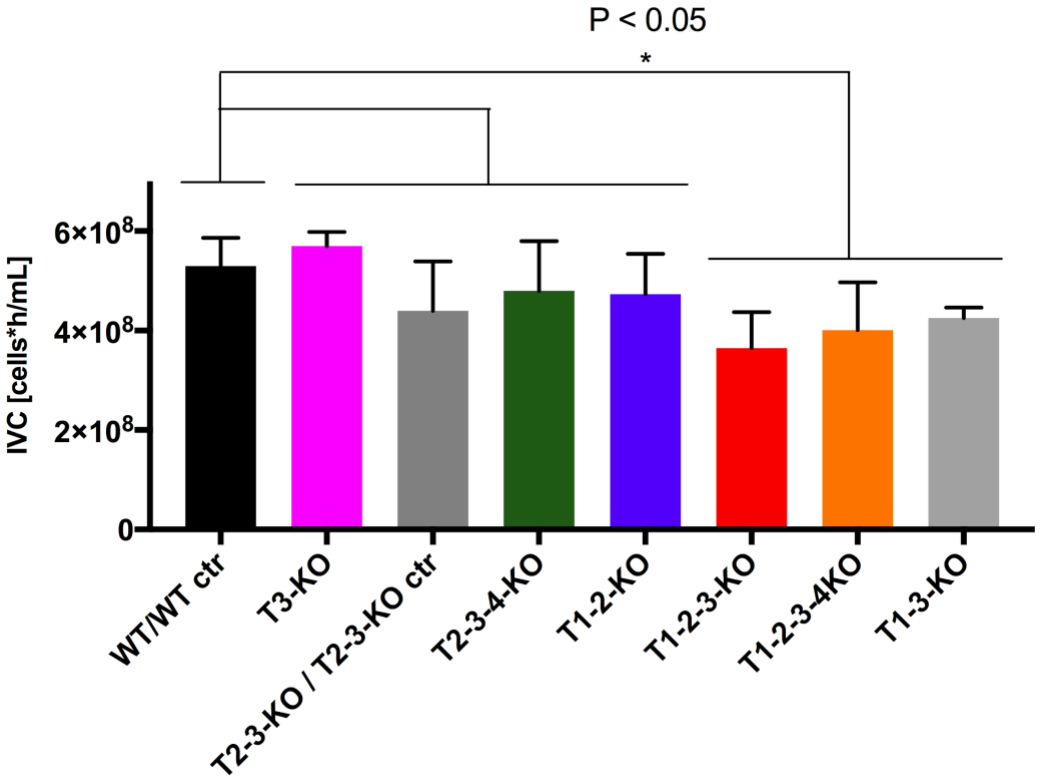
Supplementary Fig. 1. Detailed HPLC histograms of analyzed N-glycan samples from parental and generated KO cell lines. Structures from MS-annotated peaks are indicated with N-glycan cartoons and represent N-glycans from total secreted cell protein.



Supplementary Fig. 2: Rituximab N-glycan profiles and cell growth after transient transfection. N-glycan analysis of rituximab purified from different B4Gal-T-KO clones and CHO-S-WT. Rituximab was harvested three days after transient transfection for purification, N-glycan labeling and analysis. Annotated N-glycan structures point at corresponding histogram peaks. VCD and cell viabilities are presented in bottom right panel.



Supplementary Fig. 3: EPO N-glycan profiles and cell growth after transient transfection. N-glycan analysis of EPO purified from different B4Gal-T-KO clones and CHO-S-WT. EPO was harvested three days after transient transfection for purification, N-glycan labeling and analysis. Annotated N-glycan structures point at corresponding histogram peaks. VCD and cell viabilities are indicated in bottom right panel.



Supplementary Fig. 4: IVC comparison of different cell lines during batch experiment. Two major groups of genotypes with (T1-2-3-KO, T1-2-3-4-KO, T1-3-KO) and without combinatorial disruption of B4Gal-T1 and –T3 are compared to IVC of WT and WT ctr clone.

Supplementary Table S1: sgRNA target sequences. The bases in red mark the PAM site.

Target gene	Target sequence (5' → 3')
B4GalT1	TATCCCATTTCGCAACCGGCGG
B4GalT2	GGAGCACCACCTACGCTATGG
B4GalT3	TGATGCTCGCGGGCACGATGG
B4GalT4	TGGAGGCCGGTATCACCCTGG

Supplementary Table S2: Oligos for sgRNA expression vector cloning.

Oligo name	Oligo sequence (5' → 3')
gRNA_B4galt1_NW_003615120.1_283507_fwd	GGAAAGGACGAAACACCTATCCCATTTCGCAACCGGCGTTTTAGAGCTAGAAAT
gRNA_B4galt1_NW_003615120.1_283507_rev	CTAAAACGCCGGTTGCGAAATGGGATAGGTGTTTCGTCCTTTCCACAAGATAT
gRNA_B4galt2_NW_003613906.1_163085_fwd	GGAAAGGACGAAACACCGGGAGCACCACCTACGCTATGTTTTAGAGCTAGAAAT
gRNA_B4galt2_NW_003613906.1_163085_rev	CTAAAACATAGCGTAGGTGGTGCTCCCGGTGTTTCGTCCTTTCCACAAGATAT
gRNA_B4galt3_NW_003614301.1_233251_fwd	GGAAAGGACGAAACACCGTGATGCTCGCGGGCACGATGTTTTAGAGCTAGAAAT
gRNA_B4galt3_NW_003614301.1_233251_rev	CTAAAACATCGTGCCCGCGAGCATCACGGTGTTTCGTCCTTTCCACAAGATAT
gRNA_B4galt4_NW_003614660.1+444821_fwd	GGAAAGGACGAAACACCGGAGGCCGGTATCACCCTGGTTTTAGAGCTAGAAAT
gRNA_B4galt4_NW_003614660.1+444767_rev	CTAAAACAAGTGAGGTCCGGCTCAAAGGTGTTTCGTCCTTTCCACAAGATAT

Supplementary Table S3: Cas9_2A_GFP and sgRNA plasmid ratios for transfections.

	Amount of transfected plasmid [µg]					
Transfection #	GFP_2A_Cas9	sgRNA T1	sgRNA T2	sgRNA T3	sgRNA T4	pmaxGFP®
1.1	1.9	0.6	0.6	0.6		
1.2	1.9	0.9		0.9		
1.3						3.8
2.1	1.9	1.9				
2.2	1.9	0.9			0.9	
2.3						3.8

Supplementary Table S4: Primer list for deep sequencing (MiSeq). The primers contain overhang sequences compatible with Illumina Nextera XT indexing (forward primer overhang: TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG, reverse primer overhang: GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG).

MiSeq_B4galt1_NW_003615120.1_283507_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGACGCATTTGTGTACCCGAGT
MiSeq_B4galt1_NW_003615120.1_283507_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAGATGGGCGGTCGTTATTCC
MiSeq_B4galt2_NW_003613906.1_163085_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAGCAGACACATCCACAGGTG
MiSeq_B4galt2_NW_003613906.1_163085_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTCCTGGGAGGCCGTTATACA
MiSeq_B4galt3_NW_003614301.1_233251_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTGGATTTCTGTATGGGGCCA
MiSeq_B4galt3_NW_003614301.1_233251_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGTGGGTCCTGTGTCGGTATC
MiSeq_B4galt4_NW_003614660.1+444821_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAGCCCCTCTTTCGACATGTG
MiSeq_B4galt4_NW_003614660.1+444821_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTGCCATAGTCCAGTTGCTGC

round of transfection	clone name	transfected with sgRNA against target	Index	Target	Reads	WT	Indel	Indel [%]	peak1 [bp]	peak2 [bp]	peak3 [bp]	peak4 [bp]	peak1 [%]	peak2 [%]	peak3 [%]	peak4 [%]
#1		untransfected ctr	1	B4galt1	344	336	8	2.33%								
			1	B4galt2	115	112	3	2.61%								
			1	B4galt3	260	254	6	2.31%								
#1	T2-3-KO	B4Gal-T1,-T2,-T3	232	B4galt1	9,711	9,670	41	0.42%								
			232	B4galt2	6,827	28	6,799	99.59%	1				99.49%			
			232	B4galt3	9,896	28	9,868	99.72%	-5				99.18%			
#1	T1-2-3-KO A	B4Gal-T1,-T2,-T3	233	B4galt1	7749	74	7675	99.05%	1				98.66%			
			233	B4galt2	5679	10	5669	99.82%	-2	-1			49.13%	50.54%		
			233	B4GalT3	10404	76	10328	99.27%		1			99.18%			
#1		B4Gal-T1,-T2,-T3	234	B4galt1	1,316	1,275	41	3.12%								
			234	B4galt2	3,558	3,537	21	0.59%								
			234	B4galt3	7,169	7,111	58	0.81%								
#1	T3 KO A	B4Gal-T1,-T2,-T3	235	B4galt1	8,884	8,839	45	0.51%								
			235	B4galt2	4,928	4,903	25	0.51%								
			235	B4galt3	8,501	56	8,445	99.34%	1				99.15%			
#1		B4Gal-T1,-T2,-T3	236	B4galt1	3,425	3,395	30	0.88%								
			236	B4galt2	7,021	6,998	23	0.33%								
			236	B4galt3	9,657	9,606	51	0.53%								
#1		B4Gal-T1,-T2,-T3	237	B4galt1	180	169	11	6.11%								
			237	B4galt2	9,222	9,173	49	0.53%								
			237	B4galt3	30,326	61	30,265	99.80%	-11	-10	2	3 (inframe)	62.74%	5.44%	6.17%	25.12%
#1	WT ctr	B4Gal-T1,-T2,-T3	238	B4galt1	717	684	33	4.60%								
			238	B4galt2	10,950	10,908	42	0.38%								
			238	B4galt3	15,960	15,879	81	0.51%								
#1		B4Gal-T1,-T2,-T3	239	B4galt1	379	334	45	11.87%	1				6.33%			
			239	B4galt2	3,853	3,682	171	4.44%								
			239	B4galt3	7,629	7,576	53	0.69%								
#1		B4Gal-T1,-T2,-T3	240	B4galt1	138	126	12	8.70%								
			240	B4galt2	5,459	18	5,441	99.67%	-18 (inframe)				99.01%			
			240	B4galt3	12,117	57	12,060	99.53%	-2	1			49.67%	49.43%		
#1		B4Gal-T1,-T2,-T3	241	B4galt1	11,981	100	11,881	99.17%	1				98.98%			
			241	B4galt3	8,833	53	8,780	99.40%	1				99.02%			
			241	B4galt2		not available										
#1		B4Gal-T1,-T2,-T3	242	B4galt1	6,789	6,749	40	0.59%								
			242	B4galt2	3,435	10	3,425	99.71%	-18 (inframe)				99.42%			
			242	B4galt3	5,189	310	4,879	94.03%	-29	1			55.91%	37.77%		
#1		B4Gal-T1,-T2,-T3	243	B4galt1	23,393	23,313	80	0.34%								
			243	B4galt2	11,549	11,510	39	0.34%								
			243	B4galt3	17,756	17,664	92	0.52%								
#1		B4Gal-T1,-T2,-T3	244	B4galt1	2,262	2,239	23	1.02%								
			244	B4galt2	2,585	559	2,026	78.38%	-17				77.76%			
			244	B4galt3	5,221	5,202	19	0.36%								
#1		B4Gal-T1,-T2,-T3	245	B4galt1	6,170	56	6,114	99.09%	1				98.78%			
			245	B4galt2		not available										
			245	B4galt3	5,996	42	5,954	99.30%	1				99.23%			
#1		B4Gal-T1,-T2,-T3	246	B4galt1	5,743	5,699	44	0.77%								
			246	B4galt2	5,506	5,421	85	1.54%								
			246	B4galt3	9,716	40	9,676	99.59%	-41	-30 (inframe)	-10	1	52.21%	15.22%	12.43%	19.47%
#1		B4Gal-T1,-T2,-T3	247	B4galt1	157	149	8	5.10%								
			247	B4galt2		not available										
			247	B4galt3		not available										
#1	T3 KO B	B4Gal-T1,-T3	264	B4galt1	1,337	1,323	14	1.05%								
			264	B4galt2	2,620	2,611	9	0.34%								
			264	B4galt3	4,332	78	4,254	98.20%	-10	1			33.38%	64.45%		
#1		B4Gal-T1,-T3	265	B4galt1	12,789	54	12,735	99.58%	-40				98.87%			
			265	B4galt2	4,716	4,696	20	0.42%								
			265	B4galt3	7,238	42	7,196	99.42%	1				99.05%			
#1		B4Gal-T1,-T3	266	B4galt1	6,597	59	6,538	99.11%	1				98.97%			
			266	B4galt2	4,746	4,730	16	0.34%								
			266	B4galt3	7,997	22	7,975	99.72%	-36 (inframe)				99.35%			
#1		B4Gal-T1,-T3	267	B4galt1	12,031	11,983	48	0.40%								
			267	B4galt2	4,962	4,928	34	0.69%								
			267	B4galt3	11,859	11,780	79	0.67%								
#1		B4Gal-T1,-T3	268	B4galt1	2,291	2,277	14	0.61%								
			268	B4galt2	4,019	4,011	8	0.20%								
			268	B4galt3	8,004	57	7,947	99.29%	1				99.10%			
#1	T1-3-KO	B4Gal-T1,-T3	269	B4galt1	12,603	71	12,532	99.44%	1				99.33%			
			269	B4galt2	5,910	5,886	24	0.41%								
			269	B4galt3	9,959	53	9,906	99.47%	1				99.40%			
#1		B4Gal-T1,-T3	270	B4galt1		not available										
			270	B4galt2	3,023	3,007	16	0.53%								
			270	B4galt3	8,360	8,313	47	0.56%								
#1		B4Gal-T1,-T3	271	B4galt1	3,615	2,063	1,552	42.93%	1				42.63%			
			271	B4galt2	6,617	6,584	33	0.50%								
			271	B4galt3	9,057	43	9,014	99.53%	1				99.43%			
#1		B4Gal-T1,-T3	272	B4galt1	5,133	71	5,062	98.62%	1				98.50%			
			272	B4galt2	3,571	3,555	16	0.45%								
			272	B4galt3	6,957	37	6,920	99.47%	-3 (inframe)	1			53.21%	45.92%		
#1		B4Gal-T1,-T3	273	B4galt1	966	939	27	2.80%								
			273	B4galt2	5,576	5,564	12	0.22%								
			273	B4galt3	9,673	9,619	54	0.56%								
#1	T1-2 KO A	B4Gal-T1, -T2	2	B4Gal-T3	17,602	17,519	83	0.47%								
			2	B4Gal-T4	11,489	11,452	37	0.32%								
			2	B4Galt-T2	11327	38	11289	99.66%	1				99.41%			
			2	B4Galt-T1	22211	129	22082	99.42%	1				99.25%			
#1	T1-2 KO B	B4Gal-T1, -T2	3	B4Gal-T3	26,421	26,283	138	0.52%								
			3	B4Gal-T4	18,033	17,960	73	0.40%								
			3	B4Galt-T2	23285	6	23285	99.97%	-19	-7			61.62%	37.96%		
			3	B4Galt-T1	38073	190	37883	99.50%	1				99.30%			
#1		B4Gal-T1, -T2	4	B4Gal-T3	43,719	43,498	221	0.51%								
			4	B4Gal-T4	28,523	28,380	143	0.50%								
			4	B4Galt-T1	48,166	47,866	300	0.62%								
			4	B4Gal-T2		not available										
#1		B4Gal-T1, -T2	5	B4Gal-T3	25,173	25,048	125	0.50%								
			5	B4Gal-T4	21,659	21,571	88	0.41%								
			5	B4Gal-T2	28,108	10	28,098	99.96%	-30 (inframe)	-15 (inframe)			61.44%	38.13%		
			5	B4Galt-T1	33,003	32,801	202	0.61%								
#1		B4Gal-T1, -T2	6	B4Gal-T3	10,850	10,775	75	0.69%								
			6	B4Gal-T4	9,546	9,503	43	0.45%								
			6	B4Gal-T2	11,421	11,345	76	0.67%								
			6	B4Galt-T1	15,546	15,426	120	0.77%								
#1	T1-2-KO C	B4Gal-T1, -T2	7	B4Gal-T3	21,502	21,422	80	0.37%								
			7	B4Gal-T4	17,793	17,734	59	0.33%								
			7	B4Gal-T2	17,459	41	17,418	99.77%	1				99.55%			
			7	B4Galt-T1	25,100	123	24,977	99.51%	1				99.38%			
#1		B4Gal-T1, -T2	8	B4Gal-T3	34,669	34,519	150	0.43%								
			8	B4Gal-T4	27,948	27,840	108	0.39%								
			8	B4Gal-T2	30,586	1	30,585	100.00%	-13	-7			51.59%	47.97%		
			8	B4Galt-T1		not available										
#1		B4Gal-T1, -T2	9	B4Gal-T3	80,538	80,160	378	0.47%								
			9	B4Gal-T4	71,007	70,678	329	0.46%								
			9	B4Galt-T1	106,049	463	105,586	99.56%	1				99.44%			
			9	B4Gal-T2		not available										
#1		B4Gal-T1, -T2	10	B4Gal-T3	21,637	21,538	99	0.46%								
			10	B4Gal-T4	21,151	21,051	100	0.47%								
			10	B4Galt-T1	36,998	42	36,956	99.89%	2				99.36%			
			10	B4Gal-T2		not available										
#1		B4Gal-T1, -T2	11	B4Gal-T3	20,439	20,334	105	0.51%								
			11	B4Gal-T4	20,641	20,541	100	0.48%								
			11	B4Gal-T2	18,906	18,802	104	0.55%								
			11	B4Galt-T1	29,117	29,108	209	0.71%								
#1		B4Gal-T1, -T2	12	B4Gal-T3	16,542	16,456	86	0.52%								
			12	B4Gal-T4	16,587	16,512	75	0.45%								
			12	B4Gal-T2	11,784	26	11,758	99.78%	1				99.64%			
			12	B4Galt-T1	26,473	26,291	182	0.69%								
#1		B4Gal-T1, -T2	13	B4Gal-T3	14,088	13,979	109	0.77%								
			13	B4Gal-T4	9,676	9,637	39	0.40%								
			13	B4Gal-T2	9,894	5,242	4,652	47.02%	-9 (inframe)				46.48%			
			13	B4Galt-T1	18,914	18,761	153	0.81%								
#1		B4Gal-T1, -T2	14	B4Gal-T3	7,574	7,520	54	0.71%								
			14	B4Gal-T4	4,920	4,892	28	0.57%								
			14	B4Galt-T1	10,802	10,728	74	0.69%								
			14	B4Gal-T2		not available										
#1		B4Gal-T1, -T2	15	B4Gal-T3	22,428	22,282	146	0.65%								
			15	B4Gal-T4	15,388	15,312	76	0.49%								
			15	B4Gal-T2	12,173	12,091	82	0.67%								
			15	B4Galt-T1	30,675	30,491	184	0.60%								
#1		B4Gal-T1, -T2	16	B4Gal-T3	15,400	15,303	97	0.63%								
			16	B4Gal-T4	10,578	10,536	42	0.40%								
			16	B4Gal-T2	10,995	2	10,993	99.98%	-20	5			58.36%	37.22%		
			16	B4Galt-T1	22,169	92	22,077	99.59%	1				99.40%			

Supplementary Table S5: Deep sequencing results of single cell derived clones. Detailed sequencing analysis results for B4Gal-T1, -T2, -T3 and –T4 of generated clones. Clones selected for batch analysis are marked in red with out-of frame indels in blue.

#2	T2-3-KO ctr	B4Gal-T1	17	B4Gal-T3	6,616	13	6,603	99.80%	-5				99.14%				
			17	B4Gal-T4	6,394	6,357	37	0.58%									
			17	B4Gal-T2	7,259	15	7,244	99.79%	1				99.70%				
			17	B4Gal-T1	14,052	13,798	254	1.81%									
#2	T1-2-3-KO B	B4Gal-T1	18	B4Gal-T3	7,119	12	7,107	99.83%	-5				99.27%				
			18	B4Gal-T4	9,046	9,004	42	0.46%									
			18	B4Gal-T2	9,825	21	9,804	99.79%	1				99.70%				
			18	B4Gal-T1	17,038	78	16,960	99.54%	1				99.38%				
#2		B4Gal-T1	19	B4Gal-T3	5,894	7	5,887	99.88%	-5				99.46%				
			19	B4Gal-T4	7,368	7,335	33	0.45%									
			19	B4Gal-T2	6,823	16	6,807	99.77%	1				99.62%				
			19	B4Gal-T1		not available											
#2	T1-2-3-KO C	B4Gal-T1	20	B4Gal-T3	6,273	22	6,251	99.65%	-5				99.06%				
			20	B4Gal-T4	6,999	6,971	28	0.40%									
			20	B4Gal-T2	6,154	15	6,139	99.76%	1				99.59%				
			20	B4Gal-T1	11,857	53	11,804	99.55%	1				99.33%				
#2	T1-2-3-KO D	B4Gal-T1	21	B4Gal-T3	20,773	14	20,759	99.93%	-5				99.31%				
			21	B4Gal-T4	25,645	25,538	107	0.42%									
			21	B4Gal-T2	21,571	50	21,521	99.77%	1				99.68%				
			21	B4Gal-T1	46,242	198	46,044	99.57%	1				99.38%				
#2		B4Gal-T1	22	B4Gal-T3	10,350	11	10,339	99.89%	-5				99.39%				
			22	B4Gal-T4	12,077	12,027	50	0.41%									
			22	B4Gal-T2	9,694	17	9,677	99.82%	1				99.72%				
			22	B4Gal-T1	23,000	104	22,896	99.55%	1				99.28%				
#2		B4Gal-T1	23	B4Gal-T3	16,741	12	16,729	99.93%	-5				99.38%				
			23	B4Gal-T4	19,149	19,068	81	0.42%									
			23	B4Gal-T2	20,266	37	20,229	99.82%	1				99.71%				
			23	B4Gal-T1	36,909	38	36,871	99.90%	-45 (inframe)				99.38%				
#2		B4Gal-T1	24	B4Gal-T3	21,880	16	21,864	99.93%	-5				99.29%				
			24	B4Gal-T4	27,536	27,399	137	0.50%									
			24	B4Gal-T2	22,188	51	22,137	99.77%	1				99.68%				
			24	B4Gal-T1		not available											
#2		B4Gal-T1	25	B4Gal-T3	9,682	2	9,680	99.98%	-5				99.38%				
			25	B4Gal-T4	8,062	8,013	49	0.61%									
			25	B4Gal-T2	10,632	24	10,608	99.77%	1				99.67%				
			25	B4Gal-T1	18,181	82	18,099	99.55%	1				99.26%				
#2		B4Gal-T1	26	B4Gal-T3	3,818	2	3,816	99.95%	-5				99.24%				
			26	B4Gal-T4	4,375	4,348	27	0.62%									
			26	B4Gal-T2	3,099	6	3,093	99.81%	1				99.81%				
			26	B4Gal-T1	8,138	8,082	56	0.69%									
#2		B4Gal-T1	27	B4Gal-T3	6,785	27	6,758	99.60%	-5				98.87%				
			27	B4Gal-T4	7,632	7,590	42	0.55%									
			27	B4Gal-T2	13,917	27	13,890	99.81%	1				99.71%				
			27	B4Gal-T1	10,794	10,702	92	0.85%									
#2		B4Gal-T1	28	B4Gal-T3	9,229	8	9,221	99.91%	-5				99.32%				
			28	B4Gal-T4	5,567	5,527	40	0.72%									
			28	B4Gal-T2	10,590	24	10,566	99.77%	1				99.67%				
			28	B4Gal-T1	15,927	71	15,856	99.55%	1				99.28%				
#2		B4Gal-T1	29	B4Gal-T3	6,804	2	6,802	99.97%	-5				99.49%				
			29	B4Gal-T4	6,019	5,991	28	0.47%									
			29	B4Gal-T2	6,549	13	6,536	99.80%	1				99.77%				
			29	B4Gal-T1	17,397	17	17,380	99.90%	-33 (inframe)				99.36%				
#2		B4Gal-T1	30	B4Gal-T3	5,455	13	5,442	99.76%	-5				99.03%				
			30	B4Gal-T4	8,214	8,169	45	0.55%									
			30	B4Gal-T2	10,347	21	10,326	99.80%	1				99.74%				
			30	B4Gal-T1	13,099	25	13,074	99.81%	-1				99.28%				
#2		B4Gal-T1	31	B4Gal-T3	5,234	2	5,232	99.96%	-5				99.33%				
			31	B4Gal-T4	8,222	8,177	45	0.55%									
			31	B4Gal-T2	8,149	17	8,132	99.79%	1				99.71%				
			31	B4Gal-T1	13,085	4	13,081	99.97%	-4				99.27%				
#2		B4Gal-T1	32	B4Gal-T3	4,238	5	4,233	99.88%	-5				99.27%				
			32	B4Gal-T4	6,115	6,079	36	0.59%									
			32	B4Gal-T2	5,802	9	5,793	99.84%	1				99.71%				
			32	B4Gal-T1		not available											
#2		B4Gal-T1	33	B4Gal-T3	9,788	6	9,782	99.94%	-5				99.46%				
			33	B4Gal-T4	13,585	13,527	58	0.43%									
			33	B4Gal-T2	12,780	26	12,754	99.80%	1				99.71%				
			33	B4Gal-T1	21,520	91	21,429	99.58%	1				99.34%				
#2		B4Gal-T1	34	B4Gal-T3	14,249	12	14,237	99.92%	-5				99.33%				
			34	B4Gal-T4	21,786	21,692	94	0.43%									
			34	B4Gal-T2	13,897	35	13,862	99.75%	1				99.65%				
			34	B4Gal-T1		not available											
#2		B4Gal-T1, -T4	35	B4Gal-T3	11,021	9	11,012	99.92%	-5				99.32%				
			35	B4Gal-T4	15,424	17	15,407	99.89%	-4	-1			42.30%	57.20%			
			35	B4Gal-T2	12,219	22	12,197	99.82%	1				99.75%				
			35	B4Gal-T1	24,647	19	24,628	99.92%	-5				99.35%				
#2	T1-2-3-4-KO B	B4Gal-T1, -T4	36	B4Gal-T3	14,988	8	14,980	99.95%	-5				99.36%				
			36	B4Gal-T4	14,802	53	14,749	99.64%	1				99.49%				
			36	B4Gal-T2	18,578	44	18,534	99.76%	1				99.60%				
			36	B4Gal-T1	31,441	141	31,300	99.55%	1				99.36%				
#2		B4Gal-T1, -T4	37	B4Gal-T3	8,911	1	8,910	99.99%	-5				99.46%				
			37	B4Gal-T4	6,664	10	6,654	99.85%	-1				99.41%				
			37	B4Gal-T2	9,828	27	9,801	99.73%	1				99.63%				
			37	B4Gal-T1	15,775	59	15,716	99.63%	-1				99.14%				
#2		B4Gal-T1, -T4	38	B4Gal-T3	4,805	0	4,805	100.00%	-5				99.40%				
			38	B4Gal-T4	3,835	15	3,820	99.61%	1				98.96%				
			38	B4Gal-T2	5,415	14	5,401	99.74%	1				99.70%				
			38	B4Gal-T1	8,577	8,515	62	0.72%									
#2		B4Gal-T1, -T4	39	B4Gal-T3	4,624	1	4,623	99.98%	-5				99.22%				
			39	B4Gal-T4	4,186	8	4,178	99.81%	-10	1			57.81%	41.33%			
			39	B4Gal-T2	7,656	13	7,643	99.83%	1				99.80%				
			39	B4Gal-T1	9,803	33	9,770	99.66%	-44				98.99%				
#2	T1-2-3-4-KO E	B4Gal-T1, -T4	40	B4Gal-T3	5,976	2	5,974	99.97%	-5				99.25%				
			40	B4Gal-T4	4,931	9	4,922	99.82%	-1				99.27%				
			40	B4Gal-T2	6,377	13	6,364	99.80%	1				99.67%				
			40	B4Gal-T1	11,240	34	11,206	99.70%	2				99.10%				
#2		B4Gal-T1, -T4	41	B4Gal-T3	5,020	1	5,019	99.98%	-5				99.54%				
			41	B4Gal-T4	4,957	0	4,957	100.00%	-31	-4			50.78%	48.43%			
			41	B4Gal-T2	6,849	13	6,836	99.81%	1				99.63%				
			41	B4Gal-T1	10,111	10,041	70	0.69%									
#2	T2-3-4-KO C	B4Gal-T1, -T4	42	B4Gal-T3	5,324	8	5,316	99.85%	-5				99.08%				
			42	B4Gal-T4	6,875	1	6,874	99.99%	-13				99.39%				
			42	B4Gal-T2	6,304	13	6,291	99.79%	1				99.75%				
			42	B4Gal-T1	12,910	12,830	80	0.62%									
#2		B4Gal-T1, -T4	43	B4Gal-T3	5,158	6	5,152	99.88%	-5				99.26%				
			43	B4Gal-T4	6,739	1	6,738	99.99%	-17	11			57.40%	41.82%			
			43	B4Gal-T2	7,130	12	7,118	99.83%	1				99.75%				
			43	B4Gal-T1	11,197	11,125	72	0.64%									
#2		B4Gal-T1, -T4	44	B4Gal-T3	3,443	9	3,434	99.74%	-5				99.22%				
			44	B4Gal-T4	5,863	6	5,857	99.90%	-19	-1			62.78%	36.65%			
			44	B4Gal-T2	4,477	13	4,464	99.71%	1				99.66%				
			44	B4Gal-T1	7,698	40	7,658	99.48%	1				99.32%				
#2		B4Gal-T1, -T4	45	B4Gal-T3	10,969	9	10,960	99.92%	-5				99.38%				
			45	B4Gal-T4	14,515	6	14,509	99.96%	-21 (inframe)				99.44%				
			45	B4Gal-T2	16,891	35	16,856	99.79%	1				99.69%				
			45	B4Gal-T1	23,912	58	23,854	99.76%	-1				99.17%				
#2		B4Gal-T1, -T4	46	B4Gal-T3	11,029	3	11,026	99.97%	-5				99.24%				
			46	B4Gal-T2	13,010	24	12,986	99.82%	1				99.74%				
			46	B4Gal-T1	25,137	117	25,020	99.53%	1				99.39%				
			46	B4Gal-T4		not available											
#2		B4Gal-T1, -T4	47	B4Gal-T3	18,731	34	18,697	99.82%	-5				99.21%				
			47	B4Gal-T4	22,467	25	22,442	99.89%	-1				99.50%				
			47	B4Gal-T2	32,734	70	32,664	99.79%	1				99.72%				
			47	B4Gal-T1	39,307	37,126	2,181	5.55%	1				5.17%				
#2		B4Gal-T1, -T4	48	B4Gal-T3	13,170	288	12,882	97.81%	-5				97.25%				
			48	B4Gal-T4	8,180	27	8,153	99.67%	-18 (inframe)	1			53.57%	45.56%			
			48	B4Gal-T2	21,176	43	21,133	99.80%	1				99.68%				
			48	B4Gal-T1	19,004	18,890	114	0.60%									
#2		B4Gal-T1, -T4	49	B4Gal-T3	3,077	1	3,076	99.97%	-5				99.35%				
			49	B4Gal-T2	7,686	18	7,668	99.77%	1				99.67%				
			49	B4Gal-T1	5,426	66	5,360	98.78%	1				98.23%				
			49	B4Gal-T4		not available											
#2		B4Gal-T1, -T4	50	B4Gal-T3	10,710	1	10,709	99.99%	-5				99.45%				
			50	B4Gal-T4	1,918	1,888	30	1.56%									
			50	B4Gal-T2	9,321	16	9,305	99.83%	1				99.76%				
			50	B4Gal-T1	21,109	20,965	144	0.68%									
#2		B4Gal-T1, -T4	51	B4Gal-T3	14,193	5	14,188	99.96%	-5				99.23%				
			51	B4Gal-T4	8,910	13	8,897	99.85%	-1				99.28%				
			51	B4Gal-T2	16,978	32	16,946	99.81%	1				99.76%				
			51	B4Gal-T1	28,652	22,884	5,768	20.13%	-13				19.43%				
#2		B4Gal-T1, -T4	52	B4Gal-T3	9,058	17	9,041	99.81%	-5				99.19%				
			52	B4Gal-T4	10,840	23	10,817	99.79%	-1	1			56.61%	42.71%			
			52	B4Gal-T2	7,321	14	7,307	99.81%	1				99.74%				
			52	B4Gal-T1	17,667	106	17,561	99.40%	1				99.11%				

#2	T1-2-3-4-KO H	B4Gal-T1, -T4	53	B4Gal-T3	7,637	13	7,624	99.83%	-5				99.25%			
			53	B4Gal-T4	8,010	14	7,996	99.83%	-1				99.38%			
			53	B4Gal-T2	8,950	27	8,923	99.70%	1				99.63%			
			53	B4Galt-T1	17,244	58	17,186	99.66%	-1				99.04%			
			54	B4Gal-T3	13,015	14	13,001	99.89%	-5				99.38%			
			54	B4Gal-T4	13,476	24	13,452	99.82%	-1				99.34%			
			54	B4Gal-T2	20,011	48	19,963	99.76%	1				99.65%			
			54	B4Galt-T1	29,738	129	29,609	99.57%	1				99.35%			
			55	B4Gal-T3	7,865	9	7,856	99.89%	-5				99.34%			
			55	B4Gal-T4	10,644	22	10,622	99.79%	-4	1			38.49%	60.76%		
			55	B4Gal-T2	9,550	26	9,524	99.73%	1				99.65%			
			55	B4Galt-T1	16,929	93	16,836	99.45%	1				99.19%			
			56	B4Gal-T3	14,803	18	14,785	99.88%	-5				99.28%			
			56	B4Gal-T4	17,190	24	17,166	99.86%	-1				99.37%			
			56	B4Gal-T2	15,842	32	15,810	99.80%	1				99.72%			
			56	B4Galt-T1	30,925	121	30,804	99.61%	1				99.43%			
			57	B4Gal-T3	15,156	20	15,136	99.87%	-5				99.33%			
			57	B4Gal-T4	18,518	28	18,490	99.85%	-1				99.46%			
			57	B4Gal-T2	18,202	52	18,150	99.71%	1				99.66%			
			57	B4Galt-T1	32,671	150	32,521	99.54%	1				99.33%			
			58	B4Gal-T3	16,846	3	16,843	99.98%	-5				99.32%			
			58	B4Gal-T4	24,595	12	24,583	99.95%	-3 (inframe)	-1			47.11%	52.43%		
			58	B4Gal-T2	16,365	34	16,331	99.79%	1				99.69%			
			58	B4Galt-T1	37,247	34	37,213	99.91%	2				99.32%			
			59	B4Gal-T3	21,909	6	21,903	99.97%	-5				99.41%			
			59	B4Gal-T4	30,936	38	30,898	99.88%	-1	1	2		37.08%	20.75%	41.80%	
			59	B4Gal-T2	29,064	58	29,006	99.80%	1				99.71%			
			59	B4Galt-T1	48,160	47,825	335	0.70%								
			60	B4Gal-T3	20,426	22	20,404	99.89%	-5				99.29%			
			60	B4Gal-T4	26,601	7	26,594	99.97%	-4	-2			50.02%	49.45%		
			60	B4Gal-T2	17,189	34	17,155	99.80%	1				99.67%			
			60	B4Galt-T1	40,628	47	40,581	99.88%	2				99.32%			
			61	B4Galt-T1	not available											
			61	B4Gal-T3	9,533	9	9,524	99.91%	-5				99.33%			
			61	B4Gal-T4	11,234	7	11,227	99.94%	-13	-1			58.23%	41.05%		
			61	B4Gal-T2	10,796	25	10,771	99.77%	1				99.67%			
			62	B4Gal-T3	5,394	6	5,388	99.89%	-5				99.24%			
			62	B4Gal-T4	6,296	4	6,292	99.94%	-17	1			83.05%	16.23%		
			62	B4Gal-T2	5,174	12	5,162	99.77%	1				99.69%			
			62	B4Galt-T1	11,855	11,756	99	0.84%								
			63	B4Gal-T3	6,171	5	6,166	99.92%	-5				99.37%			
			63	B4Gal-T4	7,930	4	7,926	99.95%	-6 (inframe)	2			61.21%	37.89%		
			63	B4Gal-T2	8,130	21	8,109	99.74%	1				99.68%			
			63	B4Galt-T1	13,880	81	13,799	99.42%	1				99.23%			
			64	B4Gal-T3	4,726	7	4,719	99.85%	-5				99.49%			
			64	B4Gal-T4	4,783	4,743	40	0.84%								
			64	B4Gal-T2	3,078	2	3,076	99.94%	1				99.90%			
			64	B4Galt-T1	11,095	11,008	87	0.78%								
			65	B4Gal-T3	5,252	2	5,250	99.96%	-5				99.56%			
			65	B4Gal-T4	6,386	7	6,379	99.89%	-2	-1			49.72%	49.87%		
			65	B4Gal-T2	5,324	15	5,309	99.72%	1				99.62%			
			65	B4Galt-T1	11,326	65	11,261	99.43%	1				99.29%			
			66	B4Gal-T3	4,888	10	4,878	99.80%	-5				99.30%			
			66	B4Gal-T4	4,969	2	4,967	99.96%	-2	37			88.61%	10.12%		
			66	B4Gal-T2	3,863	5	3,858	99.87%	1				99.82%			
			66	B4Galt-T1	13,460	19	13,441	99.86%	-2				99.28%			
			67	B4Gal-T3	4,520	7	4,513	99.85%	-5				99.25%			
			67	B4Gal-T4	5,927	8	5,919	99.87%	-18 (inframe)	-1			63.17%	36.14%		
			67	B4Gal-T2	4,521	8	4,513	99.82%	1				99.78%			
			67	B4Galt-T1	not available											
			68	B4Gal-T3	3,591	12	3,579	99.67%	-5				99.08%			
			68	B4Gal-T4	5,995	2	5,993	99.97%	-26	-3 (inframe)			64.59%	34.81%		
			68	B4Gal-T2	1,976	5	1,971	99.75%	1				99.39%			
			68	B4Galt-T1	8,228	8,165	63	0.77%								
			69	B4Gal-T3	9,287	12	9,275	99.87%	-5				99.30%			
			69	B4Gal-T4	12,445	27	12,418	99.78%	-1				99.36%			
			69	B4Gal-T2	10,084	23	10,061	99.77%	1				99.66%			
			69	B4Galt-T1	20,835	106	20,729	99.49%	1				99.38%			
			70	B4Gal-T3	7,693	7	7,686	99.91%	-5				99.34%			
			70	B4Gal-T4	12,214	18	12,196	99.85%	-5	-2	1		20.65%	37.58%	41.20%	
			70	B4Gal-T2	5,594	11	5,583	99.80%	1				99.70%			
			70	B4Galt-T1	20,763	20,619	144	0.69%								
			71	B4Gal-T3	10,896	12	10,884	99.89%	-5				99.39%			
			71	B4Gal-T4	17,203	5	17,198	99.97%	-18 (inframe)				99.48%			
			71	B4Gal-T2	11,652	23	11,629	99.80%	1				99.74%			
			71	B4Galt-T1	27,559	116	27,443	99.58%	1				99.47%			
			72	B4Gal-T3	9,530	14	9,516	99.85%	-5				99.20%			
			72	B4Gal-T4	13,414	3	13,411	99.98%	-25	33 (inframe)			71.47%	27.96%		
			72	B4Gal-T2	6,493	15	6,478	99.77%	1				99.68%			
			72	B4Galt-T1	24,048	118	23,930	99.51%	1				99.39%			
#2	T2-3-4-KO H	B4Gal-T1, -T4	73	B4Gal-T3	15,212	1	15,211	99.99%	-5				99.48%			
			73	B4Gal-T4	14,800	3	14,797	99.98%	-13				99.35%			
			73	B4Gal-T2	19,641	47	19,594	99.76%	1				99.70%			
			73	B4Galt-T1	23,493	23,335	158	0.67%								
			74	B4Gal-T3	11,643	3	11,640	99.97%	-5				99.52%			
			74	B4Gal-T2	15,097	42	15,055	99.72%	1				99.66%			
			74	B4Galt-T1												
			74	B4Gal-T4	not available											
			75	B4Gal-T3	8,845	6	8,839	99.93%	-5				99.27%			
			75	B4Gal-T4	10,519	18	10,501	99.83%	-17	-2	1		13.13%	54.29%	32.00%	
			75	B4Gal-T2	11,763	21	11,742	99.82%	1				99.72%			
			75	B4Galt-T1	15,003	39	14,964	99.74%	4				99.05%			
			76	B4Gal-T3	8,973	5	8,968	99.94%	-5				99.42%			
			76	B4Gal-T4	8,457	8,392	65	0.77%								
			76	B4Gal-T2	7,508	13	7,495	99.83%	1				99.75%			
			76	B4Galt-T1	13,542	13,439	103	0.76%								
			77	B4Gal-T3	7,170	10	7,160	99.86%	-5				99.36%			
			77	B4Gal-T4	6,109	19	6,090	99.69%	1				99.26%			
			77	B4Gal-T2	10,055	26	10,029	99.74%	1				99.55%			
			77	B4Galt-T1	12,182	6,266	5,916	48.56%	1				48.26%			
			78	B4Gal-T3	10,973	0	10,973	100.00%	-5				99.34%			
			78	B4Gal-T4	12,412	15	12,397	99.88%	-9 (inframe)	1			64.66%	31.58%		
			78	B4Gal-T2	14,138	25	14,113	99.82%	1				99.72%			
			78	B4Galt-T1	20,002	78	19,924	99.61%	1				99.47%			
			79	B4Gal-T3	7,579	12	7,567	99.84%	-5				99.33%			
			79	B4Gal-T4	8,289	8	8,281	99.90%	-2	-1			50.42%	49.16%		
			79	B4Gal-T2	7,935	18	7,917	99.77%	1				99.68%			
			79	B4Galt-T1	16,202	16,091	111	0.69%								
			80	B4Gal-T3	6,860	15	6,845	99.78%	-5				99.14%			
			80	B4Gal-T4	3,811	23	3,788	99.40%	1				98.50%			
			80	B4Gal-T2	6,669	12	6,657	99.82%	1				99.70%			
			80	B4Galt-T1	12,142	49	12,093	99.60%	1				99.42%			
			81	B4Gal-T3	14,246	31	14,215	99.78%	-5				99.24%			
			81	B4Gal-T4	14,716	17	14,699	99.88%	-4	-1			33.05%	66.35%		
			81	B4Gal-T2	17,933	51	17,882	99.72%	1				99.64%			
			81	B4Galt-T1	27,459	27,251	208	0.76%								
			82	B4Gal-T3	11,850	7	11,843	99.94%	-5				99.44%			
			82	B4Gal-T4	15,485	2	15,483	99.99%	-17				99.46%			
			82	B4Gal-T2	14,899	22	14,877	99.85%	1				99.77%			
			82	B4Galt-T1	27,198	107	27,091	99.61%	1				99.42%			
			83	B4Gal-T3	16,681	10	16,671	99.94%	-5				99.40%			
			83	B4Gal-T2	24,760	55	24,705	99.78%	1				99.69%			
			83	B4Galt-T1	34,443	187	34,256	99.46%	1				99.29%			
			83	B4Gal-T4	not available											
#2	T2-3-4-KO K	B4Gal-T1, -T4	84	B4Gal-T3	26,671	1	26,670	100.00%	-5				99.41%			
			84	B4Gal-T4	30,487	40	30,447	99.87%	-1				99.49%			
			84	B4Gal-T2	31,987	68	31,919	99.79%	1				99.69%			
			84	B4Galt-T1	51,617	51,356	261	0.51%								

Supplementary Table S6: Gene-ID overview of targeted B4Gal-T-Isoforms.

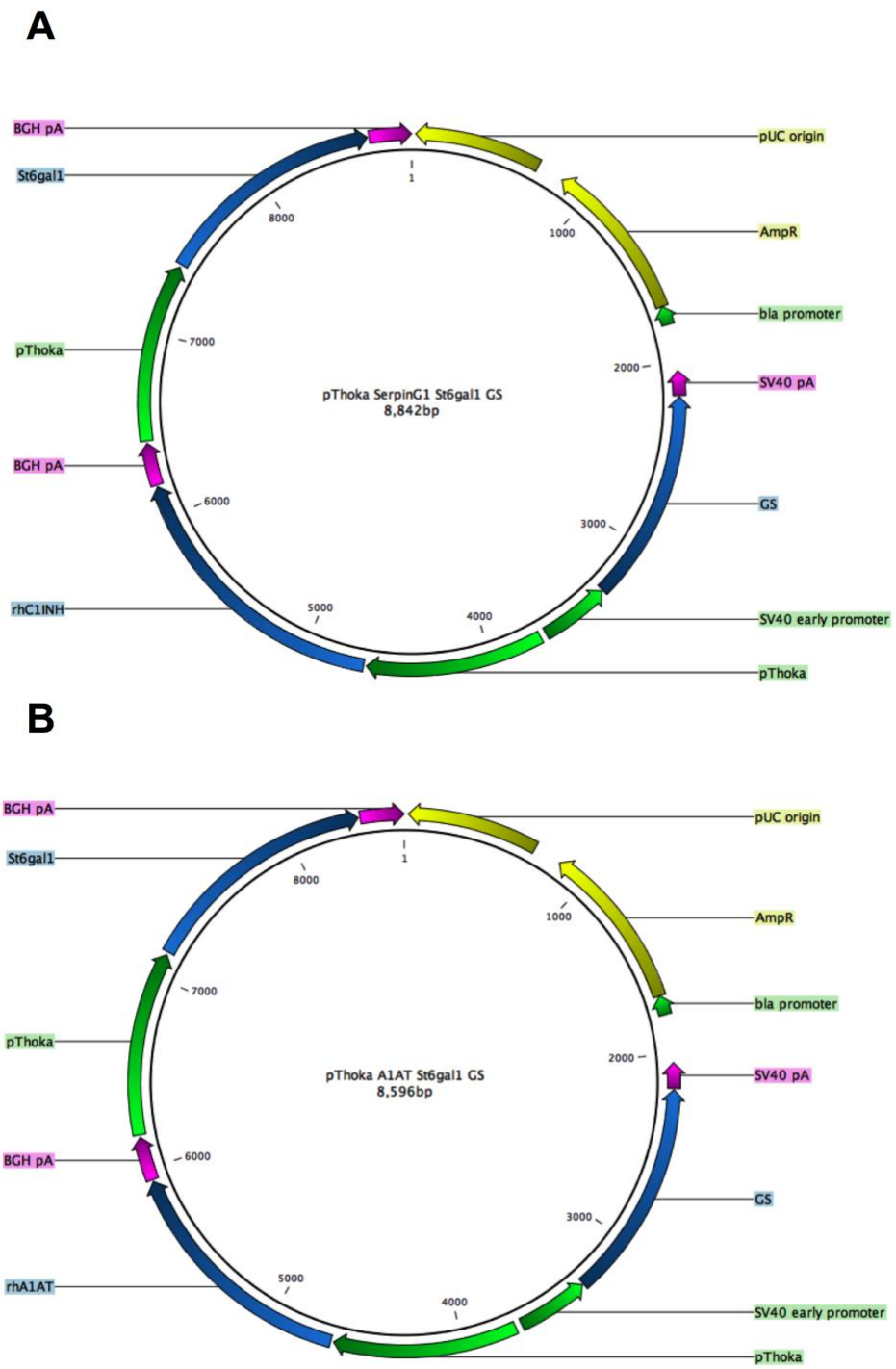
Target	**Gene-ID**
B4Gal-T1	100689430
B4Gal-T2	100689434
B4Gal-T3	100689346
B4Gal-T4	100689435

Supplementary Material for

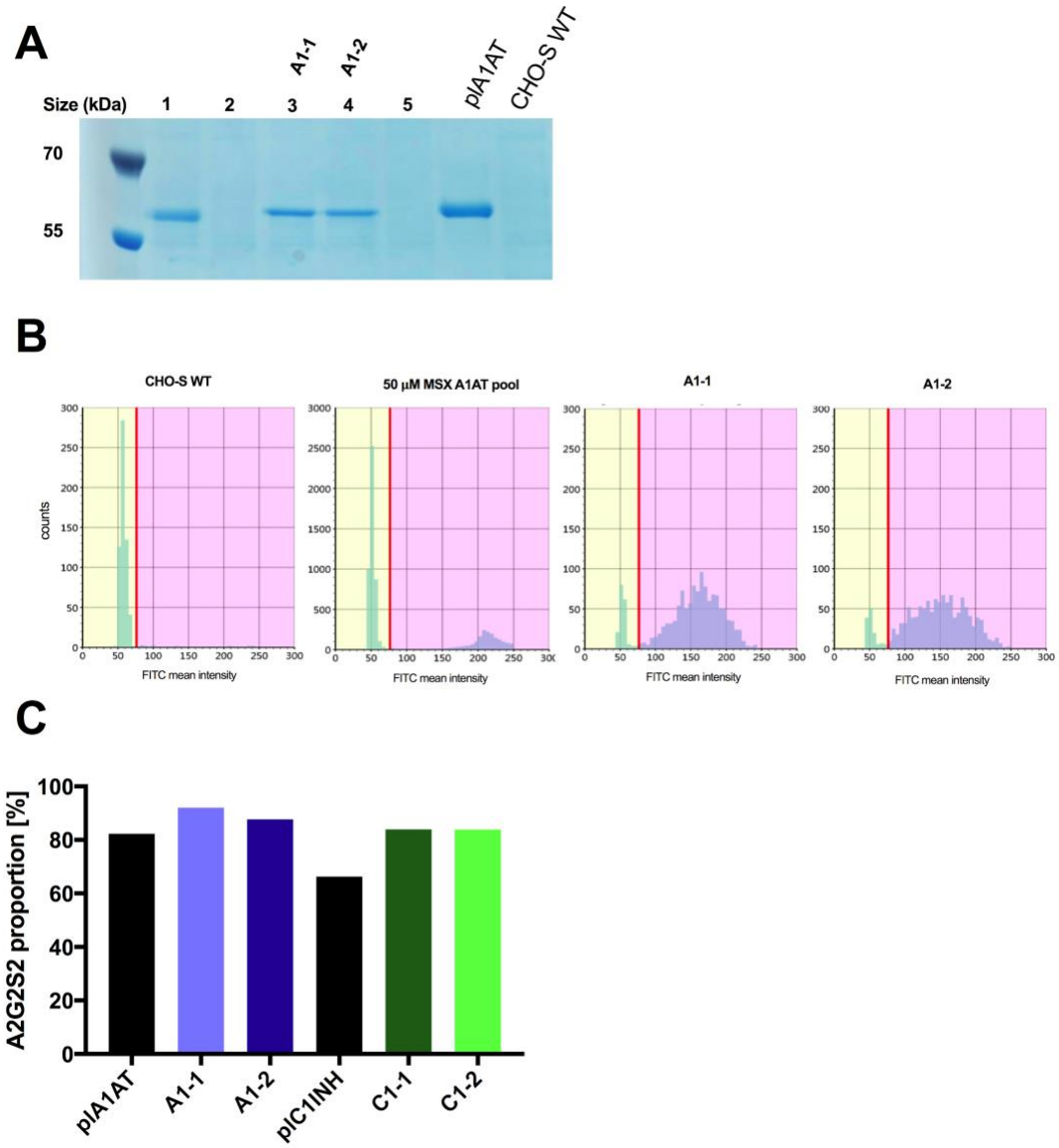
Glyco-engineered CHO cell lines producing alpha-1-antitrypsin and C1 esterase inhibitor with fully humanized N-glycosylation profiles

Thomas Amann[1], Anders Holmgaard Hansen[1], Stefan Kol[1], Henning Gram Hansen[1], Johnny Arnsdorf[1], Saranya Nallapareddy[1], Bjørn Voldborg[1], Gyun Min Lee[1,2], Mikael Rørdam Andersen[3], Helene Faustrup Kildegaard[1]

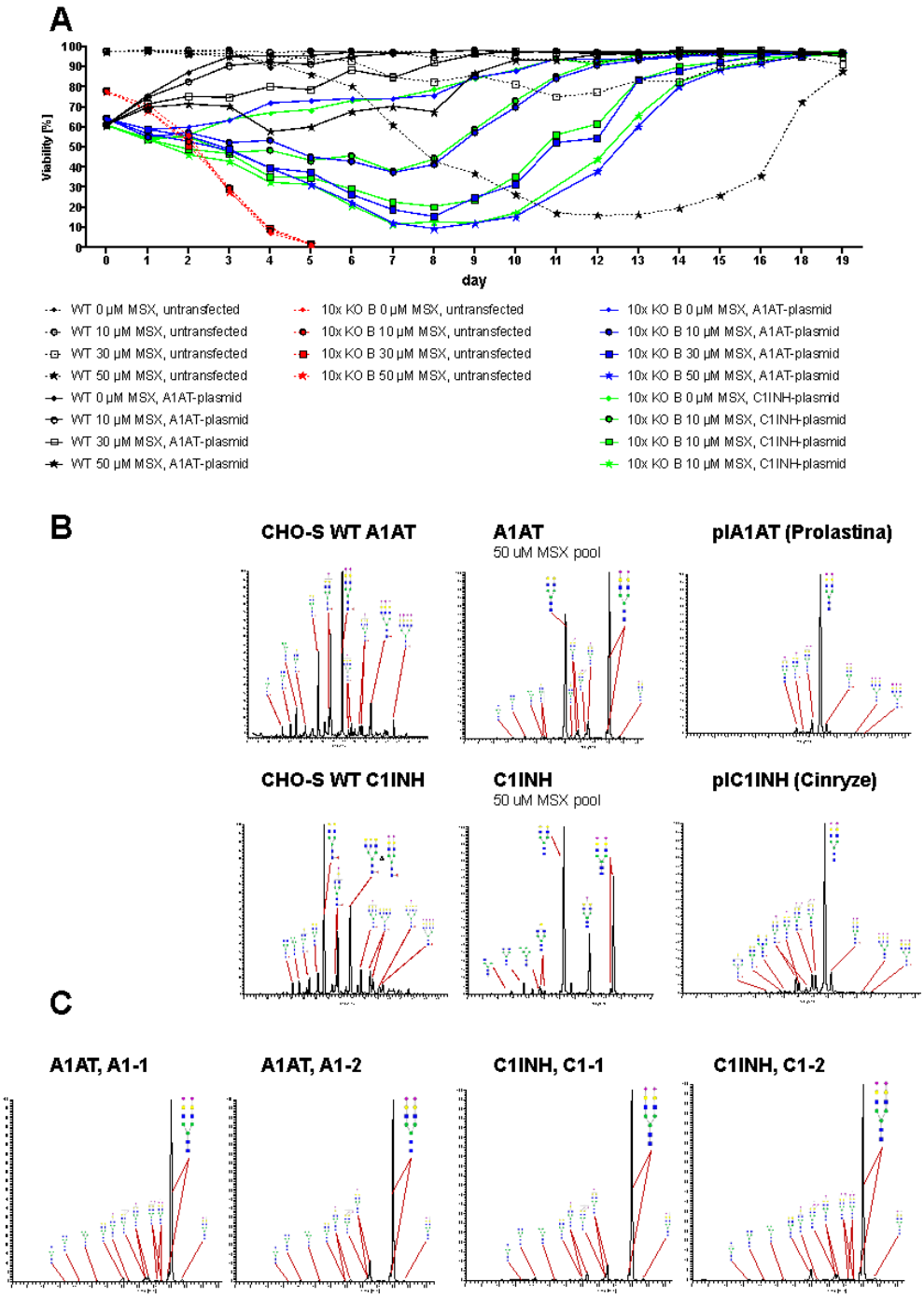
[1]The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark
[2]Department of Biological Sciences, KAIST, Daejeon, Republic of Korea
[3]Department of Biotechnology and Biomedicine, Technical University of Denmark, Kgs. Lyngby, Denmark



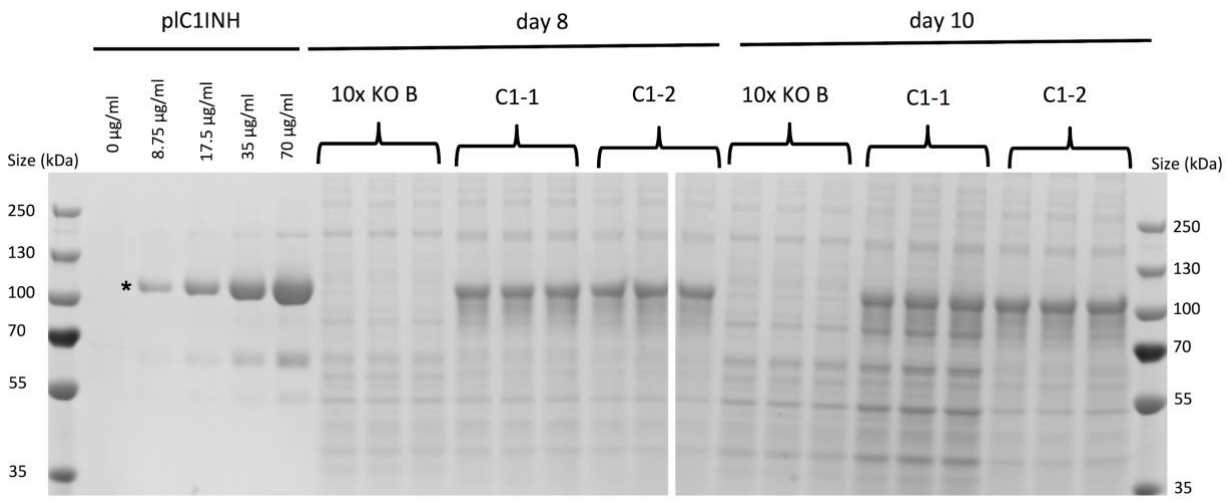
Suppl. Fig. 1: Vector design for overexpression. (A) *rh*C1INH and ST6GAL1 are under the control of mCMV-hEF-1a-5' consisting of an mCMV enhancer element, a hEF-1a promoter element, and a 5' HTLV untranslated region. The plasmid backbone consists of an ampicillin resistance cassette and SV40 early promoter in front of the GLUL (GS) sequence. Poly-A sequences (purple) terminate GLUL, *rh*A1AT and *rh*C1INH cassettes. (B) *rh*A1AT and St6gal1 are under the control of mCMV-hEF-1a-5'. The plasmid backbone consists of an ampicillin resistance cassette and SV40 early promoter in front of the GLUL (GS) sequence. Poly-A sequences (purple) terminate GLUL, *rh*A1AT and *rh*C1INH cassettes.



Suppl. Fig. 2: Screening for FITC SNA positive and *rh*A1AT producing clones. (A) Supernatant of monoclonal cell lines (clone 1 - 5) on SDS PAGE to screen for *rh*A1AT producing clones. Control samples consist of plA1AT and supernatant from non-producing CHO-S WT cells. (B) Cells of CHO-S WT, the 50 µM MSX A1AT polyclonal cell line and the clones A1-1 and A1-2 were stained with FITC-SNA. The histograms show the number of cells on y-axis and FITC mean intensity on the x-axis. CHO-S WT cells were used for gating between FITC-positive and -negative cells (red line). (C) N-glycan analysis of purified A1AT and C1INH versions. A2G2S2 proportions of purified *rh*A1AT and *rh*C1INH from different clones compared to *pl*A1AT and *pl*C1INH.



Suppl. Fig. 3: Viability of cells during MSX-based selection and N-glycan analysis of A1AT and C1INH derived from plasma as well as polyclonal and monoclonal cell lines. (A) Cell viability of non-transfected and transfected cell lines during selection with different concentrations of MSX. (B) N-glycan histograms of purified A1AT and C1INH from transiently transfected CHO-S WT and 50 µM MSX-selected polyclonal cell lines in comparison to Cinryze and Prolastin-C. (C) N-glycan analysis of purified *rh*A1AT and *rh*C1INH from four selected monoclonal producing cell lines.



Suppl. Fig. 4: SDS-PAGE gel analysis of supernatants from C1INH-producing clones and the 10x KO B parental cell line during batch cultivation. 10 µL of supernatant from day 8 and 10 of triplicate cultures as well as *pl*C1INH at different concentrations were analyzed by SDS-PAGE under reducing conditions. Intact C1INH is migrating slightly above 100 kDa as indicated by an asterisk.

Table S1: sgRNA target sequences. The bases in red mark the PAM site

Gene name of target	hypothesized KO effect	Target sequence (5'→ 3')
MGAT4A	decreased branching	GTCTACATTCGTCACTGTCGGGG
MGAT4B	decreased branching	GCTTCAGTCGCGGATCCTCTGGG
MGAT5	decreased branching	GGATGGCTACCCCCACTGCGAGG
ST3GAL3	decreased sialylation	GATCCTAGCCCACTTTCGAAAGG
ST3GAL4	decreased sialylation	GTGTCGTCGTTGTGTTGTGGTGG
ST3GAL6	decreased sialylation	GGAGTTGTGATCATTGTGAGCGG
B3GNT2	decreased elongation	GTTGGGCAAGACGCCCCCCGAGG
FUT8	no core-fucosylation	GTCAGACGCACTGACAAAGTGGG
SPPL3	hyper-glycosylation	AGAGAGACGGACGCTCCAATAGG
GLUL*	Gln-dependent growth	TCCCAAATCAGCAAACAGACTGG

*the GLUL sgRNA efficiency during KO-generation of the presented sequence was very low compared to other target sgRNAs and we recommend the usage of a different design

Table S2: Oligos for sgRNA expression vector cloning.

Oligo Name	Oligo sequence (5' → 3')
gRNA_MGAT4A_411545_fwd	GGAAAGGACGAAACACCGTCTACATTCGTCACTGTCGGTTTTAGAGCTAGAAAT
gRNA_MGAT4A_411545_rev	CTAAAACGACAGTGACGAATGTAGACCGGTGTTTCGTCCTTTCCACAAGATAT
gRNA_MGAT4B_1280368_fwd	GGAAAGGACGAAACACCGCTTCAGTCGCGGATCCTCTGTTTTAGAGCTAGAAAT
gRNA_MGAT4B_1280368_rev	CTAAAACAGAGGATCCGCGACTGAAGCGGTGTTTCGTCCTTTCCACAAGATAT
gRNA_MGAT5_327084_fwd	GGAAAGGACGAAACACCGGATGGCTACCCCCACTGCGGTTTTAGAGCTAGAAAT
gRNA_MGAT5_327084_rev	CTAAAACCGCAGTGGGGGTAGCCATCCGGTGTTTCGTCCTTTCCACAAGATAT
gRNA_ST3GAL3_244730_fwd	GGAAAGGACGAAACACCGATCCTAGCCCACTTTCGAAGTTTTAGAGCTAGAAAT
gRNA_ST3GAL3_244730_rev	CTAAAACTTCGAAAGTGGGCTAGGATCGGTGTTTCGTCCTTTCCACAAGATAT
gRNA_ST3GAL4_964386_fwd	GGAAAGGACGAAACACCGTGTCGTCGTTGTGTTGTGGGTTTTAGAGCTAGAAAT
gRNA_ST3GAL4_964386_rev	CTAAAACCACAACACAACGACGACACCGGTGTTTCGTCCTTTCCACAAGATAT
gRNA_ST3GAL6_1812502_fwd	GGAAAGGACGAAACACCGGAGTTGTGATCATTGTGAGGTTTTAGAGCTAGAAAT
gRNA_ST3GAL6_1812502_rev	CTAAAACCTCACAATGATCACAACTCCGGTGTTTCGTCCTTTCCACAAGATAT
gRNA_B3GNT2_1273293_fwd	GGAAAGGACGAAACACCGTTGGGCAAGACGCCCCCCGGTTTTAGAGCTAGAAAT
gRNA_B3GNT2_1273293_rev	CTAAAACCGGGGGGCGTCTTGCCCAACGGTGTTTCGTCCTTTCCACAAGATAT
gRNA_FUT8_681494_fwd	GGAAAGGACGAAACACCGTCAGACGCACTGACAAAGTGTTTTAGAGCTAGAAAT
gRNA_FUT8_681494_rev	CTAAAACACTTTGTCAGTGCGTCTGACGGTGTTTCGTCCTTTCCACAAGATAT
gRNA_SPPL3_213040_fwd	GGAAAGGACGAAACACCAGAGAGACGGACGCTCCAATGTTTTAGAGCTAGAAAT
gRNA_SPPL3_213040_rev	CTAAAACATTGGAGCGTCCGTCTCTCTGGTGTTTCGTCCTTTCCACAAGATAT
gRNA_GLUL_941540_fwd	GGAAAGGACGAAACACCGGCCCAGGGAAGCCATCGGAGTTTTAGAGCTAGAAAT
gRNA_GLUL_941540_rev	CTAAAACTCCGATGGCTTCCCTGGGCCGGTGTTTCGTCCTTTCCACAAGATAT

Table S3: Primer list for deep sequencing (MiSeq). The primers contain overhang sequences compatible with Illumina Nextera XT indexing
(forward primeroverhang: TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG,
reverse primer overhang: GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG).

primer name	sequence (5' → 3')
MiSeq_MGAT4A_411545_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGACAGACAGAAGGCAAATCTACG
MiSeq_MGAT4A_411545_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTTAACAGCTACACAGGAAGAGCA
MiSeq_MGAT4B_1280368_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGGATGGGGTGTATGGAGGT
MiSeq_MGAT4B_1280368_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTTGCAGACTGCTCTCCTTGG
MiSeq_MGAT5_327084_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCATGAATCTCATGGTTTCCTTTGT
MiSeq_MGAT5_327084_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGCTTCAAGACTCAACTCTTTCCC
MiSeq_ST3GAL3_244730_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGGAAACAGCATGGGCAAAC
MiSeq_ST3GAL3_244730_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGACTGGAATGTGGATGGTGGC
MiSeq_ST3GAL4_964386_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGACACCTGATGACCACATCGT
MiSeq_ST3GAL4_964386_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGCAGGGTCCACTTCTGGATT
MiSeq_ST3GAL6_1812502_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTCACTGTCTTACTACCCACAGGA
MiSeq_ST3GAL6_1812502_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTCCTTTCATTATATTCAAGAGCCAC
MiSeq_B3GNT2_1273293_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTCACCCACCGGAGAAACAG
MiSeq_B3GNT2_1273293_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAGAAGGCAAGCAATTCGGGA
MiSeq_FUT8_681494_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTGCCCCCATGACTAGGGATA
MiSeq_FUT8_681494_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTCTGCGTTCGAGAAGCTGAAA
MiSeq_SPPL3_213040_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCGTGGAGTAACTTACCTGCTGT
MiSeq_SPPL3_213040_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAAGTGGTGAGTGTGTCCTGT
MiSeq_GLUL_941540_fwd2	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCAACCAGCACCCCTGGTT
MiSeq_GLUL_941540_rev2	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCAGCTGCCAGTCTGTTTGC

Table S4: Indels generated in ten targeted genes by CRISPR/Cas9 multiplexing.

Multiplexing round	1					2		3		4
Gene	MGAT4A	MGAT4B	MGAT5	ST3GAL4	ST3GAL6	ST3GAL3	B3GNT2	GLUL	SPPL3	FUT8
10x KO clone A	+2	-1	+1	-5/+1	+1	+1/+2	-1	-13/-10/-2	+1	+1
10x KO clone B	+2	-1	+1	-5/+1	+1	+1/+2	-1	-13/-10/-2	+1	-7/-1

Table S5: Nucleotide sequences for overexpression vectors.

vector	
SerpinG1-plasmid	agacgtcaTGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTT
	TTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTAT
	AAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCC
	GCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTC
	CAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAAC
	CCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTA
	CAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTT
	ACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCA
	GCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAA
	ACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAA
	TCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTC
	TATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGT
	GCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGC
	GCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCA
	GTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGC
AmpR	TCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGAT
	CGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATC
	CGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTG
	CCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGC
	GAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTT
	TTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAA
	TGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGA
Bla promoter	ATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCGACGGATCGGGAG
	ATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTG
	TGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGA
	ATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGACAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACT
SV40pA	AGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACA
	AGTTCCGCGGTTAGTTTTTGTATTGGAAGGGCTGGTCGCCAGTCTCATTGAGAAGGCATGTGCGGACGATGGCTTCTGTC
	ACTGCAAAGGGGTCACAATTGGCAGAGGGGCGGCGGTCTTCAAAGTAACCTTTCTTCTCCTGGCCGACAGTCCGGGGAA
	TGCGGATGCTGGCACTGCGATTGGCGACACCAGCAGAAAAGTCGTTGATGTTGGACGTTTCGTGGAACCCAGTCAGACG
	ACGGGCATTGTCCAGGCCCCCCTTGGGATCGTAGGCTCGAATGTGGTACCGGTGCCGCTTGCTTAGTTTCTCGATGGCC
	TCCTCGATGTGCTTCAGACCATTCTCCTCCCGCATGGCCTTGGTGCTAAAGTTGGTATGGCAGCCTGCACCATTCCAGTT
	CCCAGGAATGGGCTTGGGGTCAAAGGTTGCTATTACCCCAAAGTCTTCACATACTCGATGCAAGATGAAACGGGCCACCC
	AGAGATGATCTCCCATGCGGATTCCTTCACAGGGTCCTATTTGGAATTCCCACTGGGCAGGCATGACCTCAGCATTTGTT
GLUL	CCTGTAATCTTGACCCCAGCATACAAGCAGGCGCGGTAGTGAGCCTCCACGATATCCCTGCCATAGGCTTTGTCTGCGCC
	CACACCACAGTAATACGGACCTTGGGGCCCAGGAAAGCCATTGGAAGGCCAACCAAAAGGGTGCCCATCTGTTCCCATC
	AGAGTATACTCCTGTTCCATTCCAAACCAGGGGTGCTGGTTGCTCACCATGTCCATTATCCGTTTACACGAGTGCCTTAAA
	TTGGTCTCTGCAGGCTTCCGGTTGTACTTGAAAACTTCACAGAACACCAGCTTGTTGGGATCTCTGCGGAAGGGGTCCCG
	AAACATGGCAACAGGGCTGAGATACATGTCACTGTTGGAGCCCTCAGACTGAAAGGTACTAGAGCCATCAAAATTCCACT
	CAGGTAACTCTTCTACACACTTGGGCTCACAGTCCAGGGTGCGGGTTTTGCAGCGCAGTCCTTCTCCAGTACCATCAACC
	CAGATATACATGGCTTGGACTTTCTCACCCTGGGGCAGGCACAAGTACATTTGCTTGATGTTTTTGTTCAAGTGGGAACTT
	GCTGAGGTGGCCATATCGATCGAAAATGGATATACAAGCTCCCGGGAGCTTTTTGCAAAAGCCTAGGCCTCCAAAAAAGC
SV40 Promoter	CTCCTCACTACTTCTGGAATAGCTCAGAGGCAGAGGCGGCCTCGGCCTCTGCATAAATAAAAAAAATTAGTCAGCCATGG
	GGCGGAGAATGGGCGGAACTGGGCGGAGTTAGGGGCGGGATGGGCGGAGTTAGGGGCGGGACTATGGTTGCTGACTA
	ATTGAGATGCATGCTTTGCATACTTCTGCCTGCTGGGGAGCCTGGGGACTTTCCACACCTGGTTGCTGACTAATTGAGAT
	GCATGCTTTGCATACTTCTGCCTGCTGGGGAGCCTGGGGACTTTCCACACCCTAACTGACACACATTCCACAGACGTCGC
	TCGATGTACGGGCCAGATATACGCGTAGTCAATGGGAAAAACCCATTGGAGCCAAGTACACTGACTCAATAGGGACTTTC
	CATTGGGTTTTGCCCAGTACATAAGGTCAATAGGGGGTGAGTCAACAGGAAAGTCCCATTGGAGCCAAGTACATTGAGTC
	AATAGGGACTTTCCAATGGGTTTTGCCCAGTACATAAGGTCAATGGGAGGTAAGCCAATGGGTTTTTCCCATTACTGACAT
	GTATACTGAGTCATTAGGGACTTTCCAATGGGTTTTGCCCAGTACATAAGGTCAATAGGGGTGAATCAACAGGAAAGTCCC
	ATTGGAGCCAAGTACACTGAGTCAATAGGGACTTTCCATTGGGTTTTGCCCAGTACAAAAGGTCAATAGGGGGTGAGTCA
	ATGGGTTTTTCCCATTATTGGCACATACATAAGGTCAATAGGGGTGACTAGTCAGTGGGCAGAGCGCACATCGCCCACAG
mCMV-hEF-1a-5' promoter	TCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTG
	ATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTTGCCGTGAACGTTCTT
	TTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCCTTCACGCGCCCGCCGCCCTA
	CCTGAGGCCGCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAACTGCGTCCGCCG
	TCTAGGTAAGTTTAAAGCTCAGGTCGAGACCGGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCT
	CTCCACGCTTTGCCTGACCCTGCTTGCTCAACTCTACGTCTTTGTTTCGTTTTCTGTTCTGCGCCGTTACAGATCCAAGCT
	GTGACCGGCGCCTACagtgcgatCGCCACCATGGCCAGCAGACTGACACTGCTGACCCTGCTGCTCCTCCTGCTGGCTGGA
	GACAGGGCTTCCTCCAACCCCAACGCCACCAGCAGCAGCTCCCAGGACCCTGAGTCCCTCCAGGACAGGGGAGAAGGC
	AAGGTCGCCACCACCGTCATCTCCAAAATGCTCTTCGTCGAGCCCATCCTCGAGGTCAGCTCCCTCCCCACCACAAACAG

C1INH/ SerpinG1	CACAACCAACAGCGCCACCAAGATCACCGCCAACACCACCGACGAACCCACAACCCAGCCCACCACAGAGCCTACAACA CAGCCTACCATCCAGCCTACCCAACCCACCACCCAGCTCCCTACCGACTCCCCTACCCAGCCTACCACAGGCTCCTTTTG TCCCGGACCTGTGACCCTGTGCTCCGACCTGGAGTCCCATAGCACAGAGGCTGTCCTCGGAGATGCCCTGGTGGATTTC AGCCTCAAACTCTACCACGCCTTCAGCGCCATGAAGAAGGTCGAGACCAATATGGCCTTCTCCCCCTTTAGCATCGCCAG CCTGCTCACCCAAGTCCTGCTCGGAGCCGGCGAGAATACCAAGACCAACCTGGAGAGCATCCTGTCCTACCCTAAGGAC TTCACCTGCGTCCACCAGGCCCTCAAGGGCTTTACCACCAAAGGAGTCACATCCGTCAGCCAGATCTTCCATTCCCCTGA CCTCGCCATTAGGGACACATTCGTGAACGCCTCCAGGACCCTGTACAGCAGCTCCCCTAGGGTCCTGTCCAACAACAGC GACGCCAACCTGGAGCTCATTAATACATGGGTGGCCAAGAATACAAACAACAAGATTAGCAGGCTCCTGGATAGCCTGCC TTCCGACACCAGGCTCGTGCTCCTCAATGCCATCTACCTCTCCGCCAAGTGGAAGACCACATTCGACCCCAAGAAAACAA GGATGGAGCCCTTTCACTTTAAAAATAGCGTGATCAAGGTGCCCATGATGAACAGCAAGAAGTACCCTGTCGCCCACTTC ATCGACCAGACCCTGAAGGCTAAGGTGGGACAGCTCCAACTGTCCCATAATCTGAGCCTGGTCATCCTCGTGCCTCAGAA CCTGAAGCACAGGCTGGAGGACATGGAACAGGCCCTGTCCCCCAGCGTGTTTAAGGCCATCATGGAAAAACTCGAGATG TCCAAGTTTCAACCCACCCTCCTCACCCTGCCCAGAATTAAGGTCACCACAAGCCAGGACATGCTCAGCATTATGGAGAA GCTCGAGTTCTTCGATTTCTCCTACGACCTCAACCTCTGCGGCCTGACAGAAGACCCTGACCTGCAGGTGAGCGCCATGC AGCACCAGACAGTGCTGGAGCTCACCGAGACAGGAGTGGAAGCTGCTGCCGCCTCCGCTATTTCCGTGGCCAGGACCC TCCTGGTGTTCGAGGTGCAACAACCCTTCCTGTTCGTCCTGTGGGACCAACAACACAAGTTCCCTGTGTTCATGGGCAGA GTCTACGACCCCAGAGCCTGAacacagtctCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCC
BGHpA	TTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCAT TCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGC
mCMV-hEF-1a-5' promoter	GGTGGGCTCTATGGATTAAGCTCGCGTAGTCAATGGGAAAAACCCATTGGAGCCAAGTACACTGACTCAATAGGGACTTT CCATTGGGTTTTGCCCAGTACATAAGGTCAATAGGGGGTGAGTCAACAGGAAAGTCCCATTGGAGCCAAGTACATTGAGT CAATAGGGACTTTCCAATGGGTTTTGCCCAGTACATAAGGTCAATGGGAGGTAAGCCAATGGGTTTTTCCCATTACTGACA TGTATACTGAGTCATTAGGGACTTTCCAATGGGTTTTGCCCAGTACATAAGGTCAATAGGGGTGAATCAACAGGAAAGTCC CATTGGAGCCAAGTACACTGAGTCAATAGGGACTTTCCATTGGGTTTTGCCCAGTACAAAAGGTCAATAGGGGGTGAGTC AATGGGTTTTTCCCATTATTGGCACATACATAAGGTCAATAGGGGTGACTAGTCAGTGGGCAGAGCGCACATCGCCCACA GTCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGT GATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTTGCCGTGAACGTTCT TTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCCTTCACGCGCCCGCCGCCCT ACCTGAGGCCGCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAACTGCGTCCGCC GTCTAGGTAAGTTTAAAGCTCAGGTCGAGACCGGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGC TCTCCACGCTTTGCCTGACCCTGCTTGCTCAACTCTACGTCTTTGTTTCGTTTTCTGTTCTGCGCCGTTACAGATCCAAGCT GTGACCGGCGCCTACagtgcgatCGCCACCATGATCCACACCAACCTGAAGAAGAAATTCTCCTGCTGCGTGGTGTTCC TGCTGTTCGCCGTGATCTGCGTGTGGAAAGAGAAGAAGGGCTCCTACTACGACTCCTTCAAGCTGCAGACCAAAGAA TTCCAGGTGCTGAAGTCCCTGGGCAAGCTGGCCATGGGCTCCGACTCTCAGTCCGTGTCCTCCAGCTCTACCCAGGACC CCCACAGAGGCAGACAGACCCTGGGCTCTCTGAGAGGCCTGGCCAAGGCTAAGCCTGAGGCCTCCTTCCAGGTGTGGA ACAAGGACTCCTCCAGCAAGAACCTGATCCCCCGGCTGCAGAAGATCTGGAAGAACTACCTGTCCATGAACAAGTACAAG GTGTCCTACAAGGGCCCTGGCCCTGGCATCAAGTTCTCTGCCGAGGCCCTGAGATGCCACCTGAGGGACCATGTGAACG TGTCCATGGTGGAAGTGACCGACTTCCCATTCAACACCTCCGAGTGGGAGGGCTACCTGCCCAAAGAGTCCATCCGGAC CAAGGCTGGCCCTTGGGGCAGATGTGCTGTGGTGTCCTCTGCCGGCTCCCTGAAGTCCTCTCAGCTGGGCAGAGAGATC GACGACCACGACGCCGTGCTGCGGTTTAATGGCGCCCCTACCGCCAACTTCCAGCAGGACGTGGGCACCAAGACCACC ATCCGGCTGATGAACTCCCAGCTCGTGACAACCGAGAAGCGGTTCCTGAAGGACTCCCTGTACAACGAGGGCATCCTGA
ST6GAL1	TCGTGTGGGACCCCTCCGTGTACCACTCCGACATCCCCAAGTGGTATCAGAACCCCGACTACAACTTCTTCAACAACTAC AAGACCTACCGGAAGCTGCACCCCAACCAGCCCTTCTACATCCTGAAGCCCCAGATGCCCTGGGAGCTGTGGGACATTC TGCAGGAAATCTCCCCCGAGGAAATCCAGCCCAACCCCCCTTCCTCTGGCATGCTGGGCATCATTATCATGATGACCCTG TGCGACCAGGTGGACATCTACGAGTTTCTGCCCTCCAAGAGAAAGACCGACGTGTGCTACTACTACCAGAAGTTCTTCGA CTCCGCCTGCACCATGGGCGCCTACCACCCTCTGCTGTACGAGAAGAACCTCGTGAAGCACCTGAACCAGGGCACCGAC GAGGATATCTACCTGCTGGGCAAGGCCACCCTGCCTGGCTTCAGAACCATCCACTGCTGAacacagtctCTGTGCCTTCTAGT TGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAA AATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGG
BGHpA	AGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGG
A1AT plasmid	agacgtcaTGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTT TTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTAT AAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCC GCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTC CAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAAC CCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTA CAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTT ACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCA GCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAA ACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAA TCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTC
AmpR	TATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGT GCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGC

	GCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCA
	GTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGC
	TCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGAT
	CGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATC
	CGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTG
	CCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGC
	GAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTT
	TTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAA
	TGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGA
Bla promoter	ATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCGACGGATCGGGAG
	ATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTG
	TGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGA
	ATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGACCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACT
SV40pA	AGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACA
	AGTTCCGCGGTTAGTTTTTGTATTGGAAGGGCTGGTCGCCAGTCTCATTGAGAAGGCATGTGCGGACGATGGCTTCTGTC
	ACTGCAAAGGGGTCACAATTGGCAGAGGGGCGGCGGTCTTCAAAGTAACCTTTCTTCTCCTGGCCGACAGTCCGGGGAA
	TGCGGATGCTGGCACTGCGATTGGCGACACCAGCAGAAAAGTCGTTGATGTTGGACGTTTCGTGGAACCCAGTCAGACG
GLUL	ACGGGCATTGTCCAGGCCCCCCTTGGGATCGTAGGCTCGAATGTGGTACCGGTGCCGCTTGCTTAGTTTCTCGATGGCC
	TCCTCGATGTGCTTCAGACCATTCTCCTCCCGCATGGCCTTGGTGCTAAAGTTGGTATGGCAGCCTGCACCATTCCAGTT
	CCCAGGAATGGGCTTGGGGTCAAAGGTTGCTATTACCCCAAAGTCTTCACATACTCGATGCAAGATGAAACGGGCCACCC
	AGAGATGATCTCCCATGCGGATTCCTTCACAGGGTCCTATTTGGAATTCCCACTGGGCAGGCATGACCTCAGCATTTGTT
	CCTGTAATCTTGACCCCAGCATACAAGCAGGCGCGGTAGTGAGCCTCCACGATATCCCTGCCATAGGCTTTGTCTGCGCC
	CACACCACAGTAATACGGACCTTGGGGCCCAGGAAAGCCATTGGAAGGCCAACCAAAAGGGTGCCCATCTGTTCCCATC
	AGAGTATACTCCTGTTCCATTCCAAACCAGGGGTGCTGGTTGCTCACCATGTCCATTATCCGTTTACACGAGTGCCTTAAA
	TTGGTCTCTGCAGGCTTCCGGTTGTACTTGAAAACTTCACAGAACACCAGCTTGTTGGGATCTCTGCGGAAGGGGTCCCG
	AAACATGGCAACAGGGCTGAGATACATGTCACTGTTGGAGCCCTCAGACTGAAAGGTACTAGAGCCATCAAAATTCCACT
	CAGGTAACTCTTCTACACACTTGGGCTCACAGTCCAGGGTGCGGGTTTTGCAGCGCAGTCCTTCTCCAGTACCATCAACC
	CAGATATACATGGCTTGGACTTTCTCACCCTGGGGCAGGCACAAGTACATTTGCTTGATGTTTTTGTTCAAGTGGGAACTT
	GCTGAGGTGGCCATATCGATCGAAAATGGATATACAAGCTCCCGGGAGCTTTTTGCAAAAGCCTAGGCCTCCAAAAAAGC
	CTCCTCACTACTTCTGGAATAGCTCAGAGGCAGAGGCGGCCTCGGCCTCTGCATAAATAAAAAAAATTAGTCAGCCATGG
SV40 Promoter	GGCGGAGAATGGGCGGAACTGGGCGGAGTTAGGGGCGGGATGGGCGGAGTTAGGGGCGGGACTATGGTTGCTGACTA
	ATTGAGATGCATGCTTTGCATACTTCTGCCTGCTGGGGAGCCTGGGGACTTTCCACACCCTAACTGACACACATTCCACAGACGTCGC
	GCATGCTTTGCATACTTCTGCCTGCTGGGGAGCCTGGGGACTTTCCACACCCTAACTGACACACATTCCACAG
	TCGATGTACGGGCCAGATATACGCGTAGTCAATGGGAAAAACCCATTGGAGCCAAGTACACTGACTCAATAGGGACTTTC
	CATTGGGTTTTGCCCAGTACATAAGGTCAATAGGGGGTGAGTCAACAGGAAAGTCCCATTGGAGCCAAGTACATTGAGTC
	AATAGGGACTTTCCAATGGGTTTTGCCCAGTACATAAGGTCAATGGGAGGTAAGCCAATGGGTTTTTCCCATTACTGACAT
	GTATACTGAGTCATTAGGGACTTTCCAATGGGTTTTGCCCAGTACATAAGGTCAATAGGGGTGAATCAACAGGAAAGTCCC
	ATTGGAGCCAAGTACACTGAGTCAATAGGGACTTTCCATTGGGTTTTGCCCAGTACAAAAGGTCAATAGGGGGTGAGTCA
	ATGGGTTTTTCCCATTATTGGCACATACATAAGGTCAATAGGGGTGACTAGTCAGTGGGCAGAGCGCACATCGCCCACAG
mCMV-hEF-1a-5' promoter	TCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTG
	ATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTTGCCGTGAACGTTCTT
	TTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCCTTCACGCGCCCGCCGCCCTA
	CCTGAGGCCGCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAACTGCGTCCGCCG
	TCTAGGTAAGTTTAAAGCTCAGGTCGAGACCGGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCT
	CTCCACGCTTTGCCTGACCCTGCTTGCTCAACTCTACGTCTTTGTTTCGTTTTCTGTTCTGCGCCGTTACAGATCCAAGCT
	GTGACCGGCGCCTACagtgcgatCGCCACCATGCCCAGCTCCGTGAGCTGGGGCATTCTCCTCCTCGCTGGCCTGTGCTGT
	CTGGTGCCTGTGAGCCTGGCCGAAGACCCCCAAGGAGACGCTGCTCAGAAGACAGACACATCCCACCATGACCAGGAC
	CACCCCACCTTCAATAAGATCACCCCTAACCTCGCTGAGTTTGCCTTTTCCCTCTACAGGCAACTGGCCCACCAGAGCAA
	CTCCACCAATATCTTCTTTAGCCCTGTGAGCATCGCCACAGCCTTCGCCATGCTGAGCCTGGGCACCAAGGCTGATACAC
	ATGACGAGATCCTGGAAGGACTGAACTTCAACCTGACCGAGATCCCCGAGGCCCAGATCCACGAGGGCTTCCAGGAACT
	GCTGAGGACCCTGAACCAGCCTGACAGCCAGCTCCAGCTCACCACCGGCAATGGCCTCTTCCTGAGCGAGGGCCTCAA
A1AT	GCTCGTGGATAAGTTCCTGGAAGACGTGAAGAAGCTGTACCACTCCGAAGCCTTCACAGTGAACTTTGGCGACACAGAG
	GAGGCCAAGAAGCAGATCAACGACTATGTGGAGAAGGGCACCCAGGGCAAGATCGTGGACCTCGTGAAGGAGCTGGAT
	AGGGACACCGTGTTCGCTCTCGTGAACTATATCTTCTTCAAGGGCAAGTGGGAGAGGCCCTTCGAGGTGAAAGACACAG
	AGGAAGAGGACTTCCACGTCGACCAAGTGACCACAGTCAAGGTCCCCATGATGAAGAGACTGGGCATGTTCAACATCCA
	GCATTGCAAAAAGCTGAGCAGCTGGGTGCTGCTCATGAAGTATCTCGGCAACGCCACAGCCATCTTCTTCCTGCCCGATG
	AGGGCAAGCTCCAGCATCTGGAAAACGAGCTCACCCACGACATTATCACCAAGTTTCTGGAGAACGAAGACAGGAGGAG
	CGCTAGCCTCCACCTCCCCAAACTCAGCATCACCGGCACATATGACCTGAAGTCCGTCCTCGGCCAGCTGGGCATCACA
	AAGGTCTTCTCCAACGGCGCCGACCTGAGCGGAGTCACAGAAGAGGCTCCCCTGAAGCTGAGCAAGGCTGTGCATAAGG
	CCGTGCTGACAATTGACGAGAAAGGCACAGAGGCTGCCGGAGCCATGTTCCTGGAAGCTATCCCCATGAGCATCCCCCC
	CGAGGTGAAATTCAACAAACCCTTCGTGTTCCTGATGATCGAGCAGAACACCAAGTCCCCCCTCTTCATGGGCAAGGTCG
	TGAACCCCACCCAGAAGTAAacacagtctCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCT
	TGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATT
BGHpA	CTATTCTGGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCG
	GTGGGCTCTATGGATTAAGCTCGCGTAGTCAATGGGAAAAACCCATTGGAGCCAAGTACACTGACTCAATAGGGACTTTC
	CATTGGGTTTTGCCCAGTACATAAGGTCAATAGGGGGTGAGTCAACAGGAAAGTCCCATTGGAGCCAAGTACATTGAGTC

mCMV-hEF-1a-5' promoter	AATAGGGACTTTCCAATGGGTTTTGCCCAGTACATAAGGTCAATGGGAGGTAAGCCAATGGGTTTTTCCCATTACTGACAT GTATACTGAGTCATTAGGGACTTTCCAATGGGTTTTGCCCAGTACATAAGGTCAATAGGGGTGAATCAACAGGAAAGTCCC ATTGGAGCCAAGTACACTGAGTCAATAGGGACTTTCCATTGGGTTTTGCCCAGTACAAAAGGTCAATAGGGGGTGAGTCA ATGGGTTTTTCCCATTATTGGCACATACATAAGGTCAATAGGGGTGACTAGTCAGTGGGCAGAGCGCACATCGCCCACAG TCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTG ATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTTGCCGTGAACGTTCTT TTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCCTTCACGCGCCCGCCGCCCTA CCTGAGGCCGCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAACTGCGTCCGCCG TCTAGGTAAGTTTAAAGCTCAGGTCGAGACCGGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCT CTCCACGCTTTGCCTGACCCTGCTTGCTCAACTCTACGTCTTTGTTTCGTTTTCTGTTCTGCGCCGTTACAGATCCAAGCT GTGACCGGCGCCTACagtgcgatCGCCACCATGATCCACACCAACCTGAAGAAGAAATTCTCCTGCTGCGTGCTGGTGTTCC TGCTGTTCGCCGTGATCTGCGTGTGGAAAGAGAAGAAGAAGGGCTCCTACTACGACTCCTTCAAGCTGCAGACCAAAGAA TTCCAGGTGCTGAAGTCCCTGGGCAAGCTGGCCATGGGCTCCGACTCTCAGTCCGTGTCCTCCAGCTCTACCCAGGACC CCCACAGAGGCAGACAGACCCTGGGCTCTCTGAGAGGCCTGGCCAAGGCTAAGCCTGAGGCCTCCTTCCAGGTGTGGA ACAAGGACTCCTCCAGCAAGAACCTGATCCCCCGGCTGCAGAAGATCTGGAAGAACTACCTGTCCATGAACAAGTACAAG GTGTCCTACAAGGGCCCTGGCCCTGGCATCAAGTTCTCTGCCGAGGCCCTGAGATGCCACCTGAGGGACCATGTGAACG TGTCCATGGTGGAAGTGACCGACTTCCCATTCAACACCTCCGAGTGGGAGGGCTACCTGCCCAAAGAGTCCATCCGGAC CAAGGCTGGCCCTTGGGGCAGATGTGCTGTGGTGTCCTCTGCCGGCTCCCTGAAGTCCTCTCAGCTGGGCAGAGAGATC GACGACCACGACGCCGTGCTGCGGTTTAATGGCGCCCCTACCGCCAACTTCCAGCAGGACGTGGGCACCAAGACCACC
ST6GAL1	ATCCGGCTGATGAACTCCCAGCTCGTGACAACCGAGAAGCGGTTCCTGAAGGACTCCCTGTACAACGAGGGCATCCTGA TCGTGTGGGACCCCTCCGTGTACCACTCCGACATCCCCAAGTGGTATCAGAACCCCGACTACAACTTCTTCAACAACTAC AAGACCTACCGGAAGCTGCACCCCAACCAGCCCTTCTACATCCTGAAGCCCCAGATGCCCTGGGAGCTGTGGGACATTC TGCAGGAAATCTCCCCCGAGGAAATCCAGCCCAACCCCCCTTCCTCTGGCATGCTGGGCATCATTATCATGATGACCCTG TGCGACCAGGTGGACATCTACGAGTTTCTGCCCTCCAAGAGAAAGACCGACGTGTGCTACTACTACCAGAAGTTCTTCGA CTCCGCCTGCACCATGGGCGCCTACCACCCTCTGCTGTACGAGAAGAACCTCGTGAAGCACCTGAACCAGGGCACCGAC GAGGATATCTACCTGCTGGGCAAGGCCACCCTGCCTGGCTTCAGAACCATCCACTGCTGAacacagtctCTGTGCCTTCTAGT
BGHpA	TGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAA AATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGG AGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGG