Identification of novel miRNA targets in CHO cell lines and characterization of their impact on protein N-glycosylation

Ankita Singh[1], Yuzhou Fan[1], Selgin Cakal[1], Thomas Amann[1], Anders Hansen[1], Nicole Borth[2], Gyun Min Lee[3], Helene Faustrup Kildegaard[4], and Mikael Andersen[1]

[1]Technical University of Denmark
[2]BOKU
[3]KAIST
[4]Novo Nordisk AS Gobal Research

January 8, 2022

Abstract

CHO cell lines are a workhorse for the production of pharmaceutical proteins, but show some limitations in the variability and stability of N-glycosylation profiles. One promising approach to addressing this at the required systems-level is miRNA, which can regulate a large number of genes and have predictable targets. Herein, we first identified de novo 656 potential miRNAs in the CHO genome based on a combination of literature, database searching, and miRNA sequencing. We further sequenced mRNA from the same cultures, and used a combination of mRNA-miRNA correlation analysis, target prediction and literature searches to find miRNAs potentially targeting N-glycosylation. Our ten best miRNA candidates were subjected to miRNA overexpression, knockdown, or knock-out in CHO cell lines. Out of the ten candidates, four (miR-128, miR-34c, miR-30b, and miR-449a) showed positive effects on N-glycosylation and could be applied directly for CHO cell engineering. The fact that 40% of the screened targets had a desired effect, and the prediction of 656 miRNAs illustrates the massive potential of miRNA engineering in CHO.

Yuzhou Fan[1,2,3]*, Ankita Singh[1,2,4]*§, Martin Seifert[1], Junrui Li[2,10], Louise Kjeldgaard Lyhne[2,5], Selgin Deniz Cakal[2], Gerald Klanert[2,6-7], Thomas Amann[1,2,9], Anders Holmgaard Hansen[1], Nicole Borth[6-7], Gyun Min Lee[8], Helene Faustrup Kildegaard[1,5], Mikael Rørdam Andersen[2,9]§

[1]*Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800, Kgs. Lyngby, Denmark*

[2]*Department of Biotechnology and Biomedicine, Technical University of Denmark, 2800, Kgs. Lyngby, Denmark*

[3]*Current affiliation: CMC Biologics, Vandtårnsvej 83, 2860 Søborg, Denmark*

[4]*Current affiliation: Department of Biomedicine, Aarhus University 8000 Aarhus Denmark*

[5]*Current affiliation: Novo Nordisk, Novo Nordisk Park, 2880 Måløv, Denmark*

[6]*ACIB GmbH, Austrian Center of Industrial Biotechnology , Graz, Austria,*

[7]*Department of Biotechnology, BOKU — University of Natural Resources and Life Sciences Vienna, Muthgasse 18, 1190 Vienna, Austria.*

[8]*Department of Biological Sciences, KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon 305-701, Republic of Korea.*

[9]*Current affiliation: Novozymes A/S, Bagsværd, Denmark*

[10]*Sahlgrenska Cancer Center, Zhuhai, Guangdong, China*

*Shared first authors. §Corresponding author.

Abstract

CHO cell lines are a workhorse for the production of pharmaceutical proteins, but show some limitations in the variability and stability of N-glycosylation profiles. One promising approach to addressing this at the required systems-level is miRNA, which can regulate a large number of genes and have predictable targets. Herein, we first identified*de novo* 656 potential miRNAs in the CHO genome based on a combination of literature, database searching, and miRNA sequencing. We further sequenced mRNA from the same cultures, and used a combination of mRNA-miRNA correlation analysis, target prediction and literature searches to find miRNAs potentially targeting N-glycosylation. Our ten best miRNA candidates were subjected to miRNA overexpression, knockdown, or knock-out in CHO cell lines. Out of the ten candidates, four (miR-128, miR-34c, miR-30b, and miR-449a) showed positive effects on N-glycosylation and could be applied directly for CHO cell engineering. The fact that 40% of the screened targets had a desired effect, and the prediction of 656 miRNAs illustrates the massive potential of miRNA engineering in CHO.

Key words:

CHO cell, Stable microRNA engineering, Crispr/Cas9, N-glycosylation, Differential equation, Correlation analysis, miRNA.

Introduction

Recombinant therapeutic proteins provide innovative and novel medical treatment possibilities for many difficult-to-treat diseases. Sales have been estimated to be 140 billion USD in 2013 (Walsh 2014) and are continually growing. Chinese hamster ovary (CHO) cells have, as the most predominant expression platforms for recombinant therapeutic proteins, for several decades shown high adaptability and robustness in industrial scale applications and great capability of producing proteins with complex folding and human-like N-glycosylation(Wong et al. 2010; Werner, Kopp, and Schlueter 2007). In particular, CHO cell factories with better and more controlled production of desired N-glycosylation structures are of interest as the glycosylation pattern has a major role in defining the efficacy of the produced protein. This drives a need to study, understand and engineer the glycosylation pattern produced by a given CHO cell line.

Recently developed CRISPR/Cas9 technologies facilitate precise genome editing in CHO cells with significantly reduced cost and delivery time (Ronda et al. 2014; Cong et al. 2013). Thus, better guidance and tools are now available for selecting and engineering different targets in CHO cells and enables rational design and development of CHO cell lines with much faster pace and higher throughput (J. Y. Kim, Kim, and Lee 2012; Kildegaard et al. 2013; Hackl et al. 2012; Sanchez et al. 2014; North et al. 2010; Sealover et al. 2013; Maszczak-Seneczko et al. 2011; Z. Yang et al. 2015). In tandem, glycoengineering studies have shown promising results in improving and controlling the quality of therapeutic proteins(Amann et al. 2019) . Even though CHO is considered a superior host for the production of therapeutic proteins, it does not typically have all the reactions of the human N-glycan pathway (Zhang et al. 2016).

Among the targets for CHO cell engineering, microRNAs (miRNAs) have emerged as a potent and promising target in the context of improving cells for production of recombinant therapeutic proteins at the phenotype level (Jadhav et al. 2013). This is because miRNA is able to regulate complex gene networks that control

various cellular processes (Lam et al. 2015), have predictable targets in the genome, and introduce no translational burden to the host cells. miRNAs are single-stranded small non-coding RNAs (19 to 25 nucleotides) that are found in a wide range of higher eukaryotes (Mack 2007). They are able to regulate the expressions of multiple genes at the post-transcriptional level by completely or partially complementary targeting 3'-UTR of the target mRNAs (Hackl et al. 2012; Jadhav et al. 2013). So far, miRNA's have been identified to impact cell cycle progression (Jadhav et al. 2013), apoptosis (Cimmino et al. 2005), metabolism (Gao et al. 2009; Csibi et al. 2013), cell proliferation (He et al. 2007), various cancer related pathways (such as the MAPK pathway (Masliah-Planchon, Garinet, and Pasmant 2016)) in mouse and humans, while miRNA studies in CHO cells are still in an early stage (Hackl et al. 2011, 2012; Jadhav et al. 2014; Gammell et al. 2007; Clarke et al. 2012; Strotbek et al. 2013; Hammond et al. 2012). However, as miRNAs are highly conserved across species (Pasquinelli et al. 2000), knowledge of miRNAs from other mammals in association with various phenotypic effects – especially findings from the highly prolific field of human cancer biology – may be translated into CHO cells. Furthermore, the role of miRNA in regards to N-glycan metabolism and protein quality needs to be explored more. Additionally, the field of N-glycosylation engineering in CHO cells has been extensively studied and heavily patented (Z. Yang et al. 2015), but almost no studies have investigated into miRNAs that modulate N-glycosylation in CHO cells so far. Therefore, to fill the gaps in miRNA engineering in CHO cells for industrial applications, more efforts on stable microRNA engineering for improving CHO cell phenotypes are needed.

In this study, we search for and discover miRNA candidates for stable CHO cell engineering to achieve desired N-glycosylation structures. We do this by identifying targets through a hybrid strategy of targeted literature and database searching, couple with mRNA and miRNA sequencing, target prediction and differential gene and miRNA expression analysis. Based on this, we overexpress (OE), knock-down (KD) and knock out (K/O) miRNA targets using CRISPR-based technologies and a CHO-adapted miRNA expression system, finding several miRNAs with effects on N-glycosylation.

Materials and Methods

CHO cell lines

Six CHO suspension cell lines were used: Three producer cell lines, namely CHO-EPO (erythropoietin producer), CHO-CS13-1.00 (rituximab high producer), and CHO-CS13-0.02 (rituximab low producer), and three non-producer cell lines (CHO-K1, CHO-DG44, and CHO-S) were used. The CS13-0.02 and CS13-1.00 cell lines were generated in a previous study by insertion of chimeric antibody genes into a CHO DG44 parental cell line, and then selected under different MTX concentrations (0.02 μM and 1.00 μM, respectively) (S. J. Kim et al. 1998). DG44-EPO is a MTX amplified stable EPO producer which was provided as a kind gift from Prof. Gyun Min Lee. CHO-K1 and CHO-DG44 are serum-free suspension cell lines, which have been adapted in-house from ATCC adherent CHO-K1 and CHO-DG44 cell lines. CHO-S was acquired from Thermo Fisher Scientific, Waltham, MA, USA.

The growth medium for CS13-0.02, CS13-1.00 and DG44-EPO was PowerCHO-2 chemically defined, serum-free medium (Lonza, Switzerland), with 4 mM of glutamine and 0.2% of anti-clumping agent (ThermoFisher Scientific) added. CHO-DG44 cell line was grown in the same growth medium with a 10% HT supplement (10 mM of sodium hypoxanthine and 1.6 mM of thymidine, ThermoFisher Scientific). The growth medium for CHO-K1 and CHO-S was CD CHO-2 chemically defined, serum-free medium (Lonza, Visp, Switzerland), with 8 mM of glutamine and 0.2% of anti-clumping agent (ThermoFisher Scientific) added.

Bioreactor cultivations

All cells were maintained in duplicates in their corresponding growth media in shake flasks at 37@C, 5% CO2, 150 rpm prior to the batch process in bioreactor.

During batch process, cell lines were grown in their corresponding growth medium in duplicates in 2x 1.0 L Eppendorf DASGIP bioreactor systems at 37°C and 150 rpm with an initial seeding density of 4×10^5 viable

cells/mL and a working volume of 410 mL. The pH and dissolved oxygen (DO) was monitored and controlled online during the cultivation process. The pH was set to 7.2 with a deviation of 0.02, where 2M sodium carbonate and CO_2 gas flow was used to adjust the pH. DO was kept at 50% of air saturation mixed of air, O_2 and CO_2 with a gas flow of 0.6 L/h. The bioreactor cultivations were last for 7 days. Daily sampling was carried out for monitoring cell growth (Nucleocounter NC-200, ChemoMetec A/S, Allerød, Denmark), and extracellular metabolites (BioProfile 400 chemistry analyzer, Nova Biomedical,Waltham,USA). Samplings of cells for RNA analysis (transcriptomics and miRNAs) was taken on day 4 of the cultivation.

RNA purification and next-generation sequencing

RNA was extracted 5×10^7 cells using miRNeasy Mini Kit (Qiagen, Germany) for total RNA preparation following the manufacturer's instructions. RNA integrity was evaluated and confirmed using an Agilent 2100 Bioanalyzer with total RNA nano chips (Agilent Technologies, Santa Clara, Ca, USA). RNA concentration was measured using Qubit 2.0 (ThermoFisher Scientific). Multiplexed cDNA library and miRNA library generation were carried out by BGI (China) using the same total RNA samples by the TruSeq RNA Sample Preparation Kit v2 (Illumina, Inc., San Diego, CA) and TruSeq Small RNA Sample Preparation Kit (Illumina), respectively. Next-generation sequencing of both transcriptome and miRNAs were performed by BGI using a Illumina Hiseq 4000 system for paired-end sequencing.

mRNA-seq quality control, alignment, and generation of read counts

The Chinese hamster genome published previously (Lewis et al. 2013) was downloaded from NCBI (accession number PRJNA239316). An initial quality check was carried out using FastQC v0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) in non-interactive mode. Quality trimming of the poor Illumina reads was performed based on the FAstQC evaluation using Trimmomatic v0.36 (Bolger, Lohse, and Usadel 2014), adopting a paired-end mode as a sequencing mode. Alignment of all trimmed paired-end reads to the Chinese hamster genome, was achieved using a STAR v2.5.2b aligner (Dobin et al. 2013), choosing genomeGenerate mode as a runMode to generates genome indices using a genome FASTA file and GTF files which are then used to align the mRNA reads. Files obtained after alignment were converted into a binary version using SAMtools (Li et al. 2009). To calculate read counts in the form of logCPM, all of the aligned reads as well as a GFF file of the Chinese hamster genome were fed into the HTSeq v0.8.0 (Anders, Pyl, and Huber 2015) package under intersection strict mode. Output from the HTSeq package were loaded into RStudio v0.99.1261 (https://www.rstudio.com/), with the integrated development environment running on R v3.3.3 (R Development Core Team 2003) for performing further downstream analysis.

Gene expression analysis (contrast matrix + DE analysis)

Through an examination of the logCPM values, we found that a large proportion of the genes within individual samples were unexpressed. Hence, we chose to enforce a cutoff of CPM of 1 or above in at least two samples (Law et al. 2016) as it distinguished expressed genes from unexpressed genes for most of the dataset (Supplementary Table 1). Next, the genes were normalized by using TMM algorithms (Law et al. 2016; Robinson and Oshlack 2010) and the *calcNormFcator* function of edgeR (McCarthy, Chen, and Smyth 2012). Subsequently, to establish the mean variance relationship in the data to calculate weights based on the calculated mean variance, we used the relationship function *voom*(Law et al. 2014) from the*limma (Smyth, n.d.)*package. Transformation of the RNA-Seq data for linear modeling was performed using the *lmFit* function of *limma* . Consequently, differentially expressed (DE) genes were determined with the*limma lmFit* function. Finally, differential expression analysis was performed using an empirical Bayes model (*eBayes*). Moderated t -statistics and the log-odds ratio was calculated using a design matrix and a contrast matrix. Summaries of the gene sets were laid out by various functions such as *toptable* and*volcanoplot (Su et al. 2017)* to finally determine the highly differentially expressed genes in each of the contrasts.

miRNA expression analysis

The quality of fastq files was evaluated using FASTQC version 0.11.2. Sequencing adapters were trimmed using cutadapt version 1.7.1. miRNAs were search against miRBase (Release 21) for *Cricetulus griseus* ,*Rattus norvegicus* , *Homo sapiens* and *Mus musculus*using CLC genomic workbench (version 7.5) with standard settings regarding mature length variants in order to obtaining the read counts of identified microRNAs. All of the gene-specific predicted miRNA targets of human and mouse were fetched from the miRWalk 2.0 database (Dweep, Gretz, and Sticht 2014). Gene and official gene symbols were kept as an input identifier and 3' UTR as an input parameter.

Correlation analysis

Unsupervised correlation analysis between expression levels of gene and miRNA was performed by comparing the expression value of genes and counts all the predicted miRNA. Pearson and Spearman correlation was performed using the *cor* function of R.

Generation of stable miRNA knock-out (KO) and knock down (KD) cell lines

500 bp upstream and 500 bp downstream sequences of target miRNAs were selected from the CHO genome sequences to identify sgRNA targeting sites and design sgRNAs using the ZiFiT tool (*www.zifit.partners.org/ZiFiT/*). The top sgRNA hits against the upstream and downstream sequences with the lowest off-target events were selected (Supplementary Table 2). The selected sgRNA was cloned into the px458 plasmid co-expressing eGFP (Addgene) by restriction with Bbs1 and conventional ligation. For each miRNA target, one plasmid with the upstream sgRNA (up-sgRNA) and one plasmid with the downstream sgRNA (dw-sgRNA) were generated. All plasmids were verified by sequencing and purified using NucleoBond Xtra Midi EF (Macherey-Nagel, Düren, Germany), according to manufacturer's instructions.

miRNA KD/KO was performed in CHO-S using CRISPR/Cas9. CHO-S cells were maintained as suspension culture in CD CHO2 medium (Lonza) with 8 mM of glutamine (Thermo Fisher Scientific) in 250ml shake flask with 120 rpm (Infors, USA) at 37 °C and 5% CO2. Cells were seeded at 0.5×10^6 cells/ml in 6 well plate (NUNC, Denmark) 1 day before transfection with 3 mL cell culture per well. Transfection of CHO-S cells were carried out by using 3.8 µg of total plasmid DNA (50 % up-sgRNA + 50 % dw-sgRNA) together with FreeStyleTM MAX reagent as described previously (Grav et al. 2015). Transfection with 3.8 µg pmaxGFP® vector without sgRNAs (Lonza, Basel, Switzerland) was applied as control for determining transfection efficiencies. Transfection with 3.8 µg px458 plasmid was carried out for generation of mock cell lines. Two days after transfection, cells were subject to FACS in order to perform single cell sorting of fluorescent-positive cells as reported (Grav et al. 2015). 14 days after single cell sorting, the colonies were expanded into 96-well flat-bottom plates (BD Biosciences). The wells with confluent cells were split, and replicated plates were subject to PCR analysis of genomic DNA when close to confluent. Genomic DNA was extracted from the cell pellets using QuickExtract DNA extraction solution (Epicentre Illumina, WI) (J. S. Lee et al. 2015). PCR was performed as described in (J. S. Lee et al. 2015) in order to select colonies that contains KO/KD of the targeted miRNA. PCR primers used for verifying miRNA KO/KD were designed to amplify from about 150 bp upstream of the up-sgRNA targeting sites and about 150 bp downstream of the dw-sgRNA targeting sites (Supplementary Table 3). Selected KO/KD cell lines were further expanded, tested, and found free of mycoplasma prior to storing at -180°C.

Generation stable miRNA overexpression cell pools

Seven miRNA overexpression (OE) plasmids were designed as described previously with minor modifications (Klanert et al. 2014). Each target sequence contains eGFP sequences linked with the stem-loop sequence of a miRNA with 50 bp upstream and 50 bp downstream flanking sequence from the CHO genome. Each

target sequence (Supplementary table 4) was synthesized at GenScript (China), and cloned into HindIII and BamHI restriction sites of pcDNA3.1/Hygro(+) plasmid (Addgene) to generate seven miRNA OE plasmids. All plasmids were purified using NucleoBond Xtra Midi EF (Macherey-Nagel), according to manufacturer's instructions.

CHO-S cells were maintained in 6 well plates (as described above for the miRNA KD/KO procedure) prior to transfection. Transfections with each OE plasmid DNA was done in the 6 well plates using FreeStyle^{TM} MAX reagent. Negative controls and mock cell pools were generated by transfecting cells with nucleotide-free water and pcDNA3.1/Hygro(+) based plasmid with a target sequence containing the eGFP sequence but no regulatory miRNA sequence, respectively. Transfections were carried out in duplicates with each plasmid. All cell cultures were kept in 6 well plates at 37degC, 5% CO2 after transfection. Transfection efficiency is determined by relative percentage of GFP cells in each culture comparing to the negative control 24 hours post transfection using Celigo Imaging Cell Cytometer (Nexcelom Bioscience, Lawrence, MA).

Two days after transfection, the transfected cells were transferred to a 125mL shake flask and were selected in 10mL of selection medium (CD CHO medium with 8 mM Gln, 0.2% AcA and 400µg/mL hygromycin B). During selection, cell culture was maintained at 37°C with 5% CO2 at 120 rpm, and medium was changed every 3–4 days. Cell growth was monitored using the NucleoCounter NC-200 every 2-3 days. After 14 days of selection, stable OE cell pools were generated when viability of each miRNA OE cell culture and mock cell culture are all above 94%, while viability of negative control cell culture is 0%. All stable cell pools were further expanded, tested, and found free of mycoplasma prior storing in the cell bank at -180°C.

Characterization of miRNA KO/KD cell lines or OE cell pools in batch culture

For each KO/KD or OE, two representative engineering cell lines/pools were selected and further evaluated with a 6-day 50mL batch culture in CD CHO-2 medium with 8 mM of glutamine and 0.2% of anti clumping agent, at 37@C, 150 rpm, 5% CO2 in duplicated 250 mL shake flasks. Daily sampling was carried out for monitoring cell growth (Nucleocounter NC-200), and extracellular metabolites (BioProfile 400 chemistry analyzer). 15 mL supernatant were harvested on day 4, and 10x concentrated using Amicon Ultra-15 (Merck, Darmstadt, Germany) for N-Glycoprofiling of secretome.

Quantitation of mature miRNA

5x10^6 cells were sampled for quantitation of mature miRNA on day 2 of the cultivation. RNA extraction was performed using a Trizol-based method as reported by (Klanert et al. 2014), and quantitation of mature miRNA were performed by quantitative real-time PCR (RT-qPCR) using miScript® SYBR® Green PCR Kit together with miScript Primer Assays (Supplementary Table 5). Synthesis was carried out using 5x miScript HiFlex Buffer and miScript Reverse Transcriptase Mix according to manufacturer's instructions. Reaction mixtures contained 12.5 µL 2x QuantiTect SYBR Green PCR Master Mix, 2.5 µL 10x miScript Universal Primers, 2.5 µL 10x miScript Primer Assay, 5 µL RNase-free water and 2.5 µL cDNA template. Amplification was executed with the following conditions: 95@C for 15 min; 40×: 94@C for 15 s, 55@C for 30 s and 70@C for 30s. Each PCR reaction had 4 replicates. Primer specificity was verified by a melting curve analysis of the PCR products with a temperature gradient of 0.2@C/s from 65@C to 95@C. The expression levels of each mature miRNA relative to a house-keeping miRNA (cgr-miR-185-5p) were determined using the 2–$^{\Delta\Delta^{\urcorner}T}$method(Klanert et al. 2014).

Quantitation of target gene expression

Quantitative real-time PCR (RT-qPCR) was performed on these genes using miScript® SYBR® Green PCR Kit (Integrated DNA Technologies, Coralville, IA) together with corresponding primers (Supplementary Table 6), according to manufacturer's instructions. Each primer pairs were passed linearity check (r^2 >0.98) using 4-fold serial dilutions of cDNA samples over 5 grades as well as amplification efficiency check (90% and 110%). Amplification was executed with the following conditions: 95@C for 15 min; 40×: 94@C for 15 s, 52@C for 30 s and 70@C for 30s. Each PCR reaction had 4 replicates, and every PCR plate included

template controls. Primer specificity was verified by a melting curve analysis. GAPDH was used as a housekeeping gene control and the relative fold change of the gene expression was calculated by the $2^{-\Delta\Delta^{\circ}T}$ method (Klanert et al. 2014).

N-Glycoprofiling of secreted proteins

Sample preparation for N-glycoprofiling was carried out using GlycoWorks RapiFluor-MS N-Glycan Kit (Waters, Milford, MA) according to the manufacturer's instructions. Labeled N-Glycans were analyzed by a LC-MS system using a Thermo Ultimate 3000 HPLC with fluorescence detector coupled to a Thermo Velos Pro Iontrap MS, as described previously (Grav et al. 2015). Relative amounts of the glycans was measured by integrating the areas under the normalized fluorescence spectrum peaks with Thermo Xcalibur software (Thermo Fisher Scientific, Waltham, MA).

Results

Identification of miRNAs in the Chinese hamster genome

Our first step was to identify as many miRNAs in the CHO genome as possible. However, current annotation of miRNAs in CHO genome is very much incomplete, and so, we decided to do *de novo* prediction, partially based on sequence analysis (described here) and partially on miRNA sequencing (See below). First we examined the Chinese hamster, human, and mouse genomes for predicted and identified miRNAs. In total, 656 miRNAs were found to align to the Chinese hamster, human and/or mouse genome. Of these, 277 were uniquely (sequence unique) mapping to the hamster genome, 353 to human. Further, 4 miRNAs were identical in human, mouse and hamster, 4 in human and hamster, 6 miRNA in human and mouse, and 12 in mouse and hamster. Out of the 656 miRNAs, 489 were predicted with the help of miRWalk to regulate one or more reactions of the N-glycan pathway (Supplementary Table 1).

Identification of genes in CHO protein N-glycosylation

First, we wanted to identify CHO genes in the N-glycosylation pathway as well as miRNAs putatively targeting these genes. To this end, a list of 75 genes assigned to the N-glycosylation pathway was collected from the KEGG pathway database (Kanehisa et al. 2019)). To further sophisticate this list, genes were grouped in categories such as mannosylation, galactosylation, fucosylation etc. based on literature mining. Furthermore, genes were classified on the basis of whether up- or down-regulation would be desirable in order to have mature N-glycosylation (Supplementary Table 7).

Generation of mRNA and miRNA expression profiles

In order to experimentally determine the expression patterns of CHO miRNAs and to be able to correlate these to gene expression patterns for genes of interest, we set up a mRNA/miRNA profiling experiment. We selected a panel of three non-producing cell lines and three protein-producing cell lines, and cultivated them in biological duplicates in batch bioreactors. CHO-S, CHO DG44 and CHO-K1 are all non-expressing cell lines; CS13-0.03 is an IgG low-producer, CS13-1.00 is an IgG high-producer, and DG44-EPO produces Erythropoietin (EPO). Culture performance of the cell lines that are used for transcriptomics and miRNAseq analysis are shown in Figure 1, all details are available in Supplementary Data File 1. As can be seen, replicates are consistent, but there are differences between the lines. CHO-S and DG44-EPO grow faster and grow to higher VCD than other cell lines. Glucose was depleted around day 5 for all cell lines except the DG44 cell line. CHO-K1 produce the most lactate and consume the most glucose during the growth phase (day 0-4) compared to other cell lines, indicating inefficient carbon flow from glycolysis to support cell growth. CHO-K1 and CHO-S also show relatively high glutamine consumption levels and high $NH4^{+}$

production levels along the cell culture comparing to other cell lines. This may be due to their growth medium used is different from other the cell lines.

RNA samples were taken on day 4, purified, and sent for mRNA and miRNA sequencing.

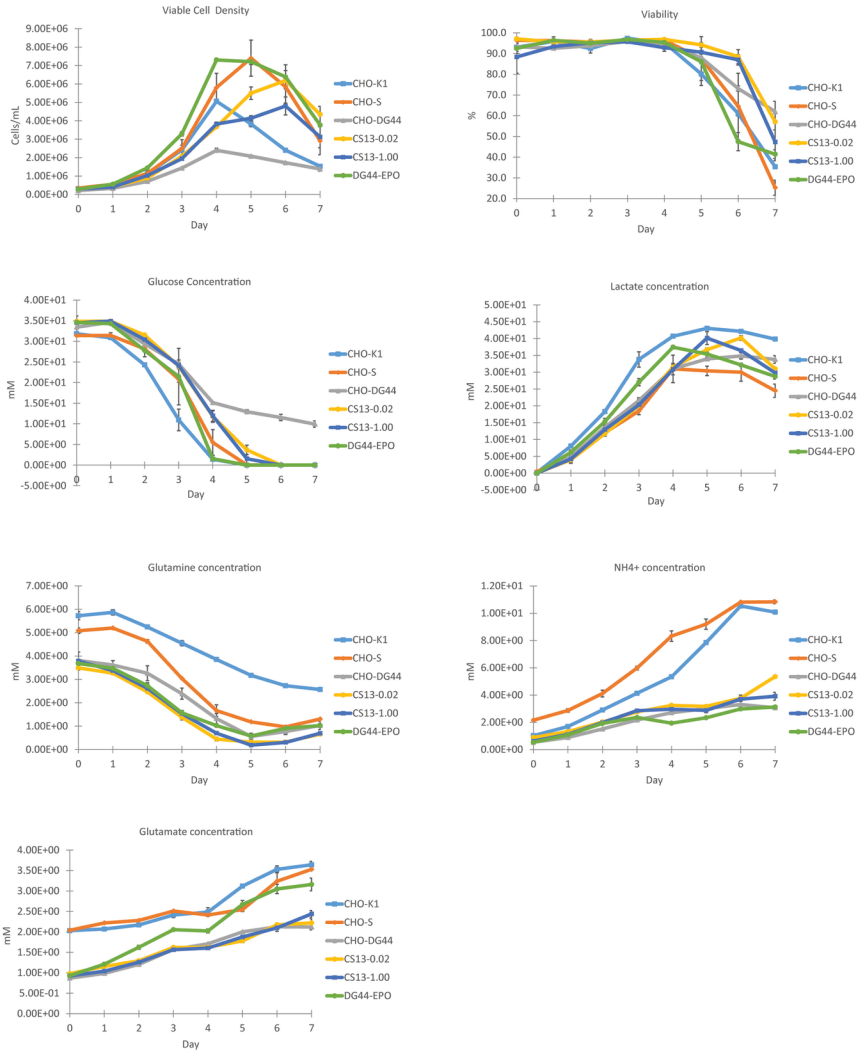


Figure 1 - Comparison of 6 cell lines used for RNAseq in bioreactor batch culture in duplicates. All RNA samples were taken on day 4 of cultivation. Error bars are showing standard deviation of duplicate culture.

RNA-seq data quality control, alignment, and normalization

All reads were subjected to quality control and trimming. Trimmed mRNA and miRNA reads were aligned to the Chinese hamster genome and read counts (logCPM) were calculated. Post-alignment, data was normalized for comparability between samples. To evaluate the normalization, library sizes of the sample sets were compared before and after normalization, providing adjusted library size pre- and post-normalization. As no problems with the aligned and normalized data were apparent, this was used for further analysis. Data on the alignment counts are available for the glycosylation genes in Supplementary Table 8, and for the mRNA in Supplementary Table 9. As a further aid, the miRNA has been sorted according to predicted gene targets.

Differential expression (DE) analysis of mRNA and miRNAs

In order to understand the general differences in the transcriptomic landscape and the miRNAs, we performed DE analysis across various combinations of producer and parental cell lines using a standard pipeline.

Since the expression profiles and N-glycan profile of the cell lines varies with the protein of interest (e.g. erythropoietin, rituximab) and the cell line, we defined four different contrasts for comparison and analyzed up- and downregulated genes involved in glycosylation as well as the set of 806 miRNAs potentially targeting glycosylation genes. Significantly differentially expressed gene sets (adj.p-value <= 0.05) were counted across various contrasts (Supplementary Table 10).

In a first comparison, we examined the effect of EPO-production by comparing the DG44 parental cell line with the DG44 EPO producer. Here, 24 glycosylation genes, and 258 glycosylation-targeting miRNA's (in the following just denoted glyco-genes and glyco-miRNAs) were found to be up-regulated, 36 glyco-genes, 106 glyco-miRNA's to be down-regulated leaving 15 glyco-genes, 442 glyco-miRNA's not significantly differentially expressed (See Supplementary Table 11 for details on genes and Supplementary Table 12 for details on miRNA).

In a similar comparison for IgG-production between the CS13-1.00 high producer and the DG44 parental cell line, 18 glyco-genes, 71 glyco-miRNA's and 28 glyco-genes, 350 glyco-miRNA's were found to be up and down-regulated respectively (genes in Supplementary Table 13, miRNA in Supplementary Table 14).

As a third comparison, we examined the difference between the CS13-1.00 high producer and the CS13-0.02 low producer cell lines were performed separately to identify traits specific to high production versus low. As expected, the significantly changed genes and miRNAs were much lower in this comparison, than the producer/non-producer comparisons above. Here, we saw 11 glyco-genes, 40 glyco-miRNA's and 11 glyco-genes, 29 glyco-miRNA's to be up- and down-regulated respectively (genes in Supplementary Table 15, miRNA in Supplementary Table 16).

As a fourth comparison aimed at identifying common traits for producer cells (irrespective of their high or low producer status), we compared all producer cell lines (i.e. DG44-EPO producer along with CS13 high and low producer) with all parental cell lines (i.e. DD44-WT, CHOS-P and CHO-K1). This comparison yielded an interesting set of 19 glyco-genes, 104 glyco-miRNA's and 37 glyco-genes, 324 glyco-miRNA's to be up- and down-regulated respectively (genes in Supplementary Table 17, miRNA in Supplementary Table 18). This is a surprisingly high number of miRNA's, more than half of the miRNAs in the cell are differentially expressed comparing producers and non-producers.

Unsupervised linking of miRNA to mRNA: Correlation analysis

While the literature study gave links to which miRNAs and genes may be linked, and the miRwalk made a sequence-based analysis, we wanted to further the analysis by incorporating an unsupervised approach to the linking of miRNA to mRNA. In general, an unsupervised statistical correlation analysis of two related datasets allows linking across datasets in an un-biased way. Hence, we employed a correlation analysis of all miRNAs in the set to miRNA's belonging to the N-glycan pathway. We chose to use two well-established correlation co-efficients: Pearson and Spearman. Pearson correlation allowed us to identify linearized relationships and Spearman correlation measures the ranked strength and direction of association between the expression variables of gene and miRNA's. Here, cell line-specific bias was minimized by averaging the expression values of each mRNA across 6 cell lines and correlating it with the expression value of its corresponding miRNA. We thus identified a list of miRNAs with a negative expression correlation by selecting only those gene-miRNA combinations that were negatively correlated (Supplementary Table 19). Similarly, positive correlations were identified. As an example, miR-6715b-5p, miR-30b-3p and miR-335-3p was found to repress eight and seven gene targets out of eight, seven and fifteen in total.

Supervised linking of miRNA to mRNA

As an addition to the unsupervised approach above, we wanted to include a supervised approach as well for linking genes to miRNAs. In order to do this, we made a correlation analysis based on the direction of statistically significant changes in the comparison of the EPO producers to the parental cells, rather than the absolute RNA-seq counts (as used in the un-supervised approached). As negative correlations, we identified miRNA targets that were either up-regulated in the DE of gene and down regulated in the DE of miRNA or vice versa. This gave us one list of miRNAs (Supplementary Table 19). The final table (Supplementary Table 19) containing all of the positive and negative correlations from both the supervised and un-supervised analysis.

Target selection

Taking the perspective that miRNA negatively inhibit and transnationally repress mRNA, we only wanted to take those targets that were negatively regulated in the supervised and/or unsupervised analysis. This gives us a filter for miRNA which supposedly works more directly on the transcript, rather than identify miRNAs targeting transcriptional repressors/activators. The list of miRNAs with negative correlations in both the supervised and unsupervised approaches can be seen in Supplementary Table 19. Combining these two types of correlation with the miRwalk analysis, the miRNA expression analysis, the miRNA DE analysis, and literature analysis, we selected a set of 10 miRNAs for modification; three for KO/KD, and seven for OE studies. The overview of the analysis is found in Table 1.

Table 1 - miRNAs engineered in this study and their hypothetical effect on CHO cell phenotypes

miRNA	**Known effect of miRNAs on various mammalian cells**	**Predicted/validated glycosylation gene targets in miRWalk2.0 database and their possible impact on N-glycosylation in human cells**	**miRNA engineering in this study**	**Reference**
miR-26b	Regulate cell proliferation and apoptosis	Gmds (8/12), Fpgt (6/12): regulate biosynthesis of GDP-fucose, and thus fucosylation	cgr-miR-26b OV	(43-46) & miRwalk 2.0
miR-30a	Regulate cell growth and productivity in CHO cells	B4galt1(8/12), SLC35A5(6/12), SLC35A3(6/12), SLC35C1(8/12).	cgr-miR-30a OV	(47, 48) & miRwalk 2.0

miR-30b	Regulate cell growth and productivity in CHO cells and changing apoptosis, O-glycosylation in human cancer cells	Fut8 (9/12): regulate fucosylation; Gale(7/12), B4galt1(8/12): regulate galactosylation; SLC35A2(7/12): regulate transport of UDP-galactose and thus galactosylation; SLC35A3(7/12): regulate transport of UDP-GlcNAc, and thus antennarity; SLC35A4(8/12), SLC35A5(6/12):): regulate transport of both UDP-galactose and UDP-GlcNAc, and thus galactosylation and antennarity; SLC35C1(8/12): regulate transport of GDP-fucose, and thus fucosylation	cgr-miR-30b OV	(47-50) & miRwalk 2.0
miR-34a	Regulate cell proliferation and apoptosis	Mgat1 (7/12), Mgat4a (10/12), Mgat4b (6/12) , Mgat5b (8/12): regulate antennarity; Fut8 (10/12): regulate fucosylation; B4galt1(8/12): regulate galactosylation	cgr-miR-34a OV	(51-54) & miRwalk 2.0
miR-34c	Regulate cell proliferation apoptosis, and DNA replication	Mgat1 (8/12), Mgat4a (10/12): regulate antennarity; Fut8 (9/12): regulate fucosylation St3gal1(8/12): regulate sialylation	cgr-miR-34c KO	(55-60) & miRwalk 2.0

miR-122		Mgat4a (9/12): regulate antennarity Fut8 (8/12): regulate fucosylation	cgr-miR-122 OV	miRwalk 2.0
miR-128	Regulate cell proliferation, apoptosis, self-renewal, and differentiation	Mgat1 (7/12), Mgat4b (8/12): regulate antennarity	cgr-miR-128 KO	(61-65) & miRwalk 2.0
miR-181a	Regulate cell proliferation, apoptosis and differentiation	B4galt1(9/12): regulate galactosylation	cgr-miR-181a and b KO in combination	(66-77) & miRwalk 2.0
miR-181b		St3gal1(8/12) : regulate sialylation B4galt1(9/12): regulate galactosylation		
miR449a	Regulate cell proliferation and apoptosis	Mgat4a (10/12): regulate antennarity Fut8 (10/12): regulate fucosylation	putative-cgr-miR449a OV	(78-81) & miRwalk 2.0

The score describes amounts of databases that predict the corresponding gene target of a miRNA out of 12 selected database in miRWalk.

These targets were examined further below.

KO/KD cell lines

In order to test the effect of reduced/loss-of-function miRNAs for three specific miRNAs – miR-128, miR-34C, and miR-181, which all have potential targets in glycosylation – and evaluate possible correlation between selected miRNAs and N-glycosylation, we set up a CRISPR-based engineering strategy.

For each target, two plasmids were generated, each expressing one specific guide RNA (gRNA) targeting the upstream and downstream region of the miRNA expression site in the CHO genome (up-sgRNA and dw-sgRNA respectively).

For each target, up-sgRNA and dw-sgRNA plasmids were co-transfected into CHO-S cells, and subsequently single cell sorting and screened. As shown in Figure 2A , KO/KD of miRNA was first verified at the molecular level by genomic DNA PCR. It was possible to isolate both miR-128, miR-34C, and miR-181 KO/KD clones which showed a clean deletion (KO). For some miR-128 and miR-34c clonal cell lines, both wt bands and KO bands were found in the gel. This indicates that the deletion may only have happened in one allele in these cell lines, thus making them KD rather than KO, and suggesting that these exist in more than one allele.

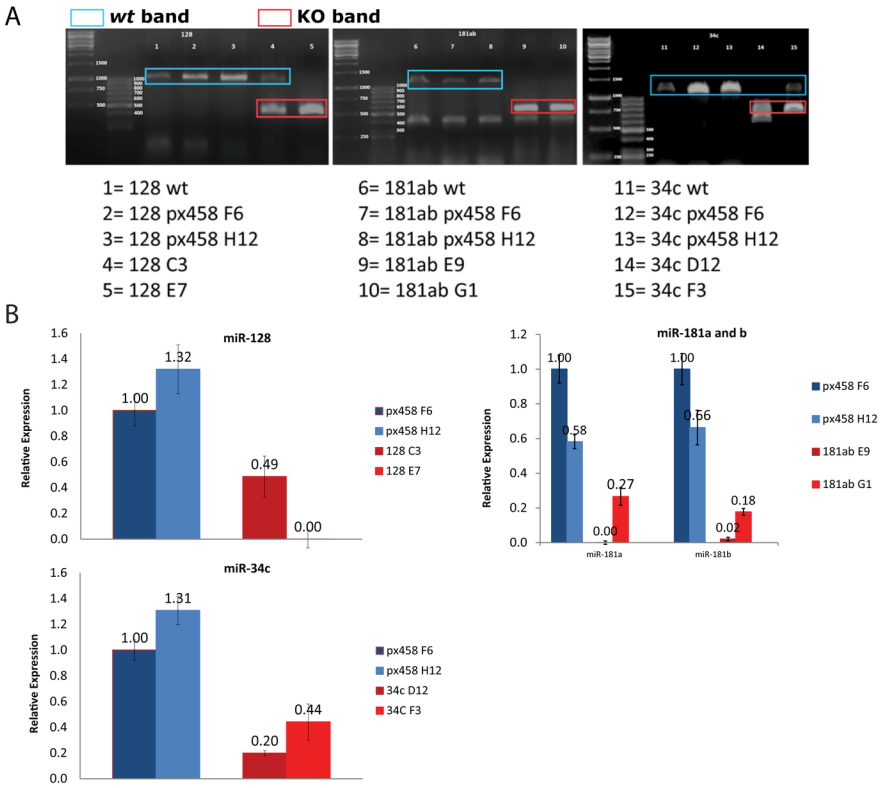


Figure 2 - Characterization of KO/KD cell lines in CHO-S. A: Out-out PCR of miRNA KO/KD in CHO-S. B: Quantification of mature miRNA by qPCR.

To further validate the miRNA KO/KD cell lines we conducted quantitative real-time PCR of corresponding miRNAs (Figure 2B). The miRNA qPCR results correlate well with the results from out-out PCR, which indicates that miR-128 E7, and miR-181ab E9 are to be considered full miR KO cell lines, but the miR-34c D12 seems to have some miR-34 expression active, despite the negative gDNA PCR result, possibly due to a copy somewhere else in the genome. It should thus only be considered a KD.

Glycoprofiling of miRNA KO/KD cell lines

To functionally characterize the miRNA KO/KD effects, we examined N-glycosylation in the clones. Batch cultures were run in shake flasks, and N-glycosylation profile of secretomes was made from sampling day 4 of the batch culture. Figure 3 shows that miR-128 and miR-34c KO/KD increases the levels of tri- and tetra-antennary structure of N-glycans of the secreted proteins in CHO cells in comparison to the mock cell lines (px458 F6 and px458 H12). On the other hand, miRNA-181 KO cell lines demonstrate no change comparing to mock cell lines. This result indicates that miR-128 and miR-34c may regulate tri- and tetra- antennary levels in N-glycans by either regulating N-acetylglucosamine (GlcNAc) transferases that is responsible for addition of the GlcNAc unit onto β 1,4 (Mgat4a, Mgat4b), and β 1,6 branches (Mgat5), or by regulating UPD-GlcNAc transporters (SLC35A3, SLC35A4, SLC35A5), which is responsible of transporting UPD-GlcNAc from cytosol into Golgi apparatus, as a building block for further GlcNAc transferase reactions.

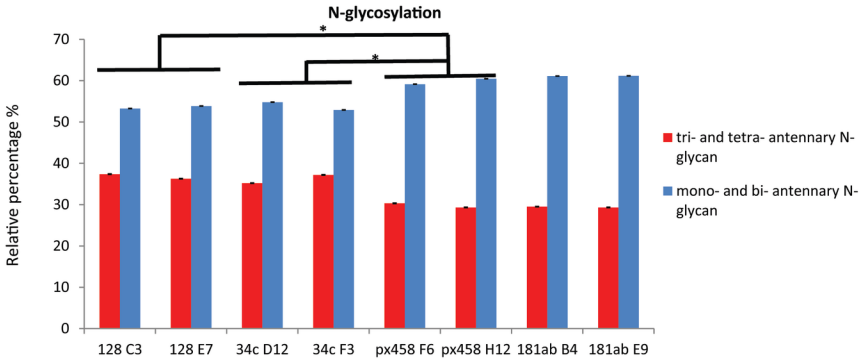


Figure 3 - N-glycan characterization of miRNA KO/KD cell lines. Statistically significant difference (*, $p<0.05$) antennary structures among secreted proteins from the mock (px458 F6 and px458 H12) and KD/KO (miR-128 C3, miR-128 E7, miR-34c D12, and miR-34c F3) cell lines based on N-glycosylation profiling. Details are found in Supplementary Figure 1.

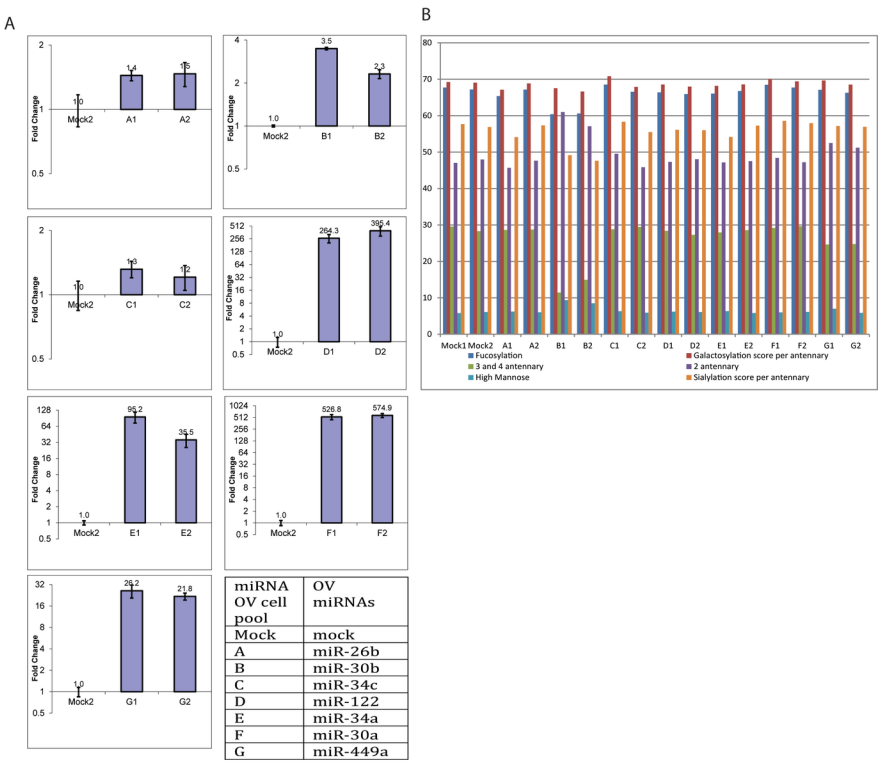
In a qPCR analysis of glycosylation gene expression in the culture (Figure 5), it seems that miR-128 KD may regulate Mgat4b and Mgat5 by increasing their expression levels. In addition, miR-34c KD may regulate the expression levels of Mgat4a and Mgat4b, and possible also Mgat5. On the other hand, SLC35A3 and SLC35A5 seems not to be directly targeted by mir-128 and miR-34c, as expression levels of SLC35A3 and SLC35A5 in the corresponding miR KD cell lines are similar to the mock cell line.

Construction and validation of miRNA over-expression (OE) cell pools

In order to test more miRNA targets, we constructed seven cell pools (Table 1), each overexpressing one miRNA. We tested the expression with qPCR, and all miRNA OE cell pools show relative overexpression of corresponding miRNAs (Figure 4A). The levels of the relative fold change are correlated well with the level of corresponding of endogenous miRNAs expressed in the original CHO-S cell lines. The miRNAs that has lower expression levels in original CHO-S cell lines, for example mir-122 and miR30a, shows higher fold change when overexpressing the corresponding miRNAs.

Glycoprofiling of OE cell pools

The next step was to explore the phenotype of the miRNA overexpression. We did this by glycoprofiling the secreted protein from the seven pools and comparing them to a mock (Figure 4B). It is clear, that only overexpressing miR-30b and miR-449a gave detectable changes in N-glycan profiles, namely reduced levels of tri- and tetra- antennary structures, and increase mono- and bi- antennary structures. This indicates that miR-30b and miR-449a may also target genes that regulate GlcNAc transferase reactions for addition GlcNac onto β 1,4 (Mgat4a, Mgat4b) and β 1,6 branches (Mgat 5), and/or target UPD-GlcNAc tansporters (SLC35A3, SLC35A4, SLC35A5), or possibly enzymes working earlier in the N-glycan pathway. According to qPCR results of possible target genes (Figure 6), we found that miR-30b levels correlate inversely with the levels of Mgat4a, SLC35A3 and SLC35A5. Thus, miR-30b seems to target Mgat4a, SLC35A3 and SLC35A5 mRNAs. Therefore, miR-30b may change antennary structures of N-glycan by simultaneously regulate GlcNAc transferase reactions and the availability of UDP-GlcNAc inside Golgi apparatus. miR-449a, on the other hand, may only target Mgat4a and reduce the levels of β 1,4 branched N-glycans.



miRNA OV cell pool	OV miRNAs
Mock	mock
A	miR-26b
B	miR-30b
C	miR-34c
D	miR-122
E	miR-34a
F	miR-30a
G	miR-449a

Figure 4. - Characterization of miRNA OE cell pools. (A) Quantification of mature miRNA by qPCR. (B) Glycoprofiling of miRNA OE cell pools (Details are found in Supplementary Table 20 and Supplementary Figure 2)

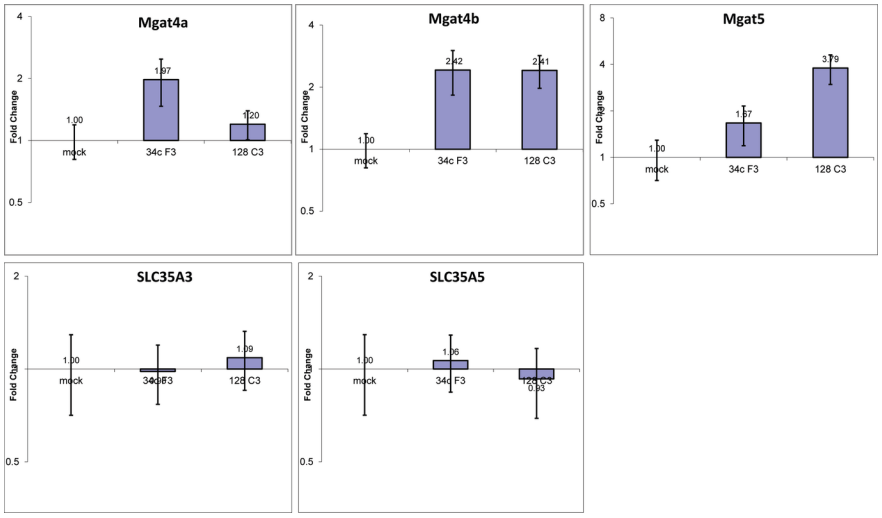


Figure 5 - qPCR analysis for possible target glycosylation genes of miR-34c and miR-128 KD cell lines.

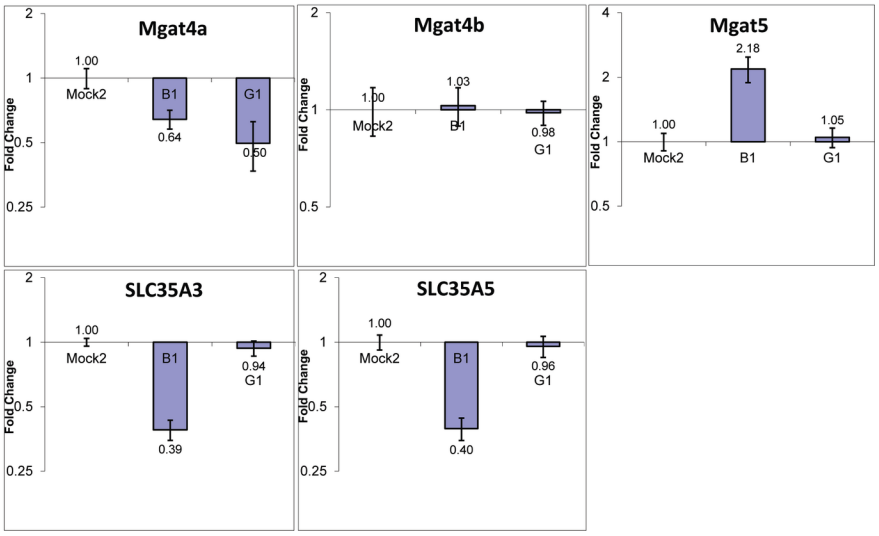


Figure 6 - qPCR analysis for possible target glycosylation genes of miR-30b OE (B1), miR-449a OE (G1) and mock cell pool.

Discussion

Along with our increased understanding of miRNAs and miRNA correlations, it has become feasible to make at least rough predictions of the targets of miRNAs. Furthermore, due to the ability of the miRNA to degrade or to silence multiple mRNAs in a highly specific and potent way, makes miRNA a highly interesting target for cell engineering; in particular of complex traits such as N-glycosylation, which involves a large number of enzymes working in concert. Hence, in this study we have tried to adopt a composite prediction and experimental strategy to closer examine on how miRNA regulating the N-glycan pathway in CHO cells

In this study, the strategy – in addition to sequence-based prediction – was to use expression data of mRNA and miRNA to link miRNA with mRNA based on correlation analysis. As we were interested in picking only genes which are inhibited by miRNA, we have filtered for only those mRNA-miRNA combinations that were negatively co-expressed. In combination with a pathway analysis, to filter for N-glycan pathways such as fucosylation, mannosylation, sialylation, galactosylation and branching of N-glycan, we could focus on relevant genes. In the end, this combination of approaches gave a list of likely miRNA-mRNA sets. Furthermore, the experimental approach and analysis of mRNA and miRNA sequencing has generated a lot of new knowledge on which miRNAs are active in these cell lines, and provide a long list of hypotheses on co-expressed genes. In total, we saw that four out of our ten miRNA targets had an effect on N-glycosylation in the tested conditions, showing that this field has an enormous potential for future applications.

For instance, in our combined analysis, we saw that miR-128 should theoretically combine with Alg9 and Man2a1, based on that mRNA expression of Alg9 and Man2a1 was down-regulated and miR-128 was up-regulated in EPO-producing cell lines as compared to parental cell lines. Suggesting knock-out of miR-128 in EPO producing cell lines would ideally increase the expression level of Alg9 and Man2a1, increasing alpha 1,3 mannosyltransferase and mannosidase alpha class2 member1 activities. However, this was not the phenotype that was seen directly, but instead, we saw a shift towards more mature (tri- and tetra-antennary) N-glycans. This fitted well with the miRwalk predictions, predicting binding to Mgat1 and Mgat4b (Glc-NAc transferases, involved in increasing antennarity). A similar effect was found for miR-34c. Here, Alg1 and B4galt2 were found in combination with miR-34c, where expression levels supported an inverse correlation, and we noticed that even though expression level of miR-34c was down regulated; it was significantly expressed in EPO producing cell lines suggesting complete functional knock-out of miR-34c to

see increase in the activity of alpha 1,3 mannosyltransferase and beta1,4 galactosyltransferase. Furthermore, Mgat1 and Mgat4 and St3gal1 (sialyltransferases) were predicted by miRwalk to be targets (Table 1). In this cases as well, KD of the miRNA increased the maturation. It thus seems - based on the prediction of the functional targets - that increasing availability of early N-glycosylation pathway enzymes improves maturation of the N-glycans. This is very important for proteins such as EPO, as mature antennae and sialylation is critical for serum half-life of EPO(Tsuda et al. 1990). Similarly, for antibodies, mannose residues can increase antibody clearance rates (Kaufman et al. 1991; K. H. Lee et al. 2019; Fabian 2010; J. Yang et al. 2015). Further, sialylation of native and recombinant EPO increases its*in vitro* biological activity, possibly by increasing its affinity to receptor on target cells(Kaufman et al. 1991). It was previously shown that making more sites available for adding sialic acid by overexpressing B4gal1, also in tandem with St3gal1(Kaufman et al. 1991; K. H. Lee et al. 2019). Terminal N-linked glycosylation is directly related to N-linked sialylation. Thus, the prediction that miR-128 and miR-34c inhibits early maturation enzymes and Mgat1 and Mgat4, fits well with increased number of antennae, and will be beneficial for certain recombinant protein production, e.g. EPO production.

For the overexpression studies, only miR30b and miR-449a out of the seven tested miRNAs gave a phenotype, again increased maturation of the N-glycans. This was interesting, as these miRNAs did not have predicted targets in the correlation analysis, but were predicted by miRwalk to have a large number of gene targets in glycosylation.

In general, our study also shows that it is hard to predict the phenotype of a miRNA perturbation solely based on sequence analysis, even when supplemented with miRNA- and mRNA-seq. One reason for this may be that miRNAs show as little complementarity as 2-8nt to their gene targets, making them hard to identify using sequence alignment. Furthermore previous studies(Kaufman et al. 1991; K. H. Lee et al. 2019; Fabian 2010) have shown that higher complementarity with the mRNA with their targets degrade the transcripts and lower complementarity leads to translational repression , thus making both the prediction more complicated, and the regulation even more complex. Even so, we managed to find a desirable phenotype in 40% of our predicted targets. Due to limited capacity, we were only able to test 10 targets, but the full data set (Supplementary Table 19) contains even more interesting future targets, for examples, miRNAs with high correlation with FUT8 fucosyltransferase, galactotransferases B4galT1-8, sialylases St3gal1-6, and sialic acid transporters SLC35A1-A5.

Conclusion

We have adopted a bifurcated comparative approach using both computational and experimental elements to identify miRNAs with an impact on N-glycan maturation. In total, 656 miRNAs were predicted, ten were experimentally characterized and four were found to have beneficial phenotypes for N-glycosylation. The present study is one among few studies aiming to assess the impact of overall miRNAs on N-glycan pathway of various CHO cell lines. The data generated from miRNA- and mRNA-seq is an asset for the scientific community as it provides the valuable understanding of some of the useful expression targets of the genes. Finally, identification of CHO-specific miRNAs presented in the study might serve as a tool to improve cell line engineering effort to increase the overall productive robustness of the cell lines.

Author contributions

Yuzhou Fan – design overall studies for both wet lab and dry lab, developed protocols, conduct experiments, coordinate with RNAseq experiment, process miRNA dry-lab data, analyze wet lab data, write manuscript.

Ankita Singh – design methods for data analysis, conduct RNAseq data analysis and miRNA-mRNA correlation data analysis, write manuscript.

Martin Seifert – conduct experiment to generate KO cell lines

Junrui Li1 – conduct bioreactor experiment to generate RNAs for sequencing

Louise Kjeldgaard Lyhne - conduct experiment to generate OV cell pools and characterized OV cell pools

Selgin Deniz Cakal - conduct experiment to generate OV cell pools and characterized OV cell pools

Gerald Klanert – support with original protocol development for miRNA OV and KO studies

Thomas Amann - support and conduct N-glycosylation sample preparation.

Anders Holmgaard Hansen - support with N-glycoprofiling by running LC-MS and provide raw data files.

Nicole Borth – support with methods and experience of miRNA engineering in general

Gyun Min Lee – supported with cell lines

Helene Faustrup Kildegaard – overall design and support to miRNA KO study and support with cell lines, manuscript editing

Mikael Rørdam Andersen – overall design and support for whole study, manuscript writing.

Refernces

Amann, Thomas, Anders Holmgaard Hansen, Stefan Kol, Henning Gram Hansen, Johnny Arnsdorf, Saranya Nallapareddy, Bjørn Voldborg, Gyun Min Lee, Mikael Rørdam Andersen, and Helene Faustrup Kildegaard. 2019. "Glyco-Engineered CHO Cell Lines Producing Alpha-1-Antitrypsin and C1 Esterase Inhibitor with Fully Humanized N-Glycosylation Profiles." *Metabolic Engineering*52 (March): 143–52.

Anders, Simon, Paul Theodor Pyl, and Wolfgang Huber. 2015. "HTSeq–a Python Framework to Work with High-Throughput Sequencing Data." *Bioinformatics* 31 (2): 166–69.

Bolger, Anthony M., Marc Lohse, and Bjoern Usadel. 2014. "Trimmomatic: A Flexible Trimmer for Illumina Sequence Data." *Bioinformatics* 30 (15): 2114–20.

Cimmino, Amelia, George Adrian Calin, Muller Fabbri, Marilena V. Iorio, Manuela Ferracin, Masayoshi Shimizu, Sylwia E. Wojcik, et al. 2005. "miR-15 and miR-16 Induce Apoptosis by Targeting BCL2." *Proceedings of the National Academy of Sciences of the United States of America* 102 (39): 13944–49.

Clarke, Colin, Michael Henry, Padraig Doolan, Shane Kelly, Sinead Aherne, Noelia Sanchez, Paul Kelly, et al. 2012. "Integrated miRNA, mRNA and Protein Expression Analysis Reveals the Role of Post-Transcriptional Regulation in Controlling CHO Cell Growth Rate." *BMC Genomics* 13 (November): 656.

Cong, Le, F. Ann Ran, David Cox, Shuailiang Lin, Robert Barretto, Naomi Habib, Patrick D. Hsu, et al. 2013. "Multiplex Genome Engineering Using CRISPR/Cas Systems." *Science* 339 (6121): 819–23.

Csibi, Alfred, Sarah-Maria Fendt, Chenggang Li, George Poulogiannis, Andrew Y. Choo, Douglas J. Chapski, Seung Min Jeong, et al. 2013. "The mTORC1 Pathway Stimulates Glutamine Metabolism and Cell Proliferation by Repressing SIRT4." *Cell* 153 (4): 840–54.

Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. 2013. "STAR: Ultrafast Universal RNA-Seq Aligner." *Bioinformatics* 29 (1): 15–21.

Dweep, Harsh, Norbert Gretz, and Carsten Sticht. 2014. "miRWalk Database for miRNA–Target Interactions." In *Methods in Molecular Biology*, 289–305.

Fabian, Dirk. 2010. "I-CAMP 2010 Australia Monday June 21 Dirk Fabian Professional Networking SPIE Webcast." *SciVee*. https://doi.org/10.4016/18748.01.

Gammell, Patrick, Niall Barron, Niraj Kumar, and Martin Clynes. 2007. "Initial Identification of Low Temperature and Culture Stage Induction of miRNA Expression in Suspension CHO-K1 Cells." *Journal of Biotechnology* 130 (3): 213–18.

Gao, Ping, Irina Tchernyshyov, Tsung-Cheng Chang, Yun-Sil Lee, Kayoko Kita, Takafumi Ochi, Karen I. Zeller, et al. 2009. "C-Myc Suppression of miR-23a/b Enhances Mitochondrial Glutaminase Expression and Glutamine Metabolism." *Nature* 458 (7239): 762–65.

Grav, Lise Marie, Jae Seong Lee, Signe Gerling, Thomas Beuchert Kallehauge, Anders Holmgaard Hansen, Stefan Kol, Gyun Min Lee, Lasse Ebdrup Pedersen, and Helene Faustrup Kildegaard. 2015. "One-Step Generation of Triple Knockout CHO Cell Lines Using CRISPR/Cas9 and Fluorescent Enrichment." *Biotechnology Journal* 10 (9): 1446–56.

Hackl, Matthias, Vaibhav Jadhav, Tobias Jakobi, Oliver Rupp, Karina Brinkrolf, Alexander Goesmann, Alfred Pühler, Thomas Noll, Nicole Borth, and Johannes Grillari. 2012. "Computational Identification of microRNA Gene Loci and Precursor microRNA Sequences in CHO Cell Lines." *Journal of Biotechnology* 158 (3): 151–55.

Hackl, Matthias, Tobias Jakobi, Jochen Blom, Daniel Doppmeier, Karina Brinkrolf, Rafael Szczepanowski, Stephan H. Bernhart, et al. 2011. "Next-Generation Sequencing of the Chinese Hamster Ovary microRNA Transcriptome: Identification, Annotation and Profiling of microRNAs as Targets for Cellular Engineering." *Journal of Biotechnology* 153 (1-2): 62–75.

Hammond, Stephanie, Jeffrey C. Swanberg, Shawn W. Polson, and Kelvin H. Lee. 2012. "Profiling Conserved microRNA Expression in Recombinant CHO Cell Lines Using Illumina Sequencing." *Biotechnology and Bioengineering* 109 (6): 1371–75.

He, Lin, Xingyue He, Lee P. Lim, Elisa de Stanchina, Zhenyu Xuan, Yu Liang, Wen Xue, et al. 2007. "A microRNA Component of the p53 Tumour Suppressor Network." *Nature* 447 (7148): 1130–34.

Jadhav, Vaibhav, Matthias Hackl, Aliaksandr Druz, Smriti Shridhar, Cheng-Yu Chung, Kelley M. Heffner, David P. Kreil, et al. 2013. "CHO microRNA Engineering Is Growing up: Recent Successes and Future Challenges." *Biotechnology Advances* 31 (8): 1501–13.

Jadhav, Vaibhav, Matthias Hackl, Gerald Klanert, Juan A. Hernandez Bort, Renate Kunert, Johannes Grillari, and Nicole Borth. 2014. "Stable Overexpression of miR-17 Enhances Recombinant Protein Production of CHO Cells." *Journal of Biotechnology* 175 (April): 38–44.

Kanehisa, Minoru, Yoko Sato, Miho Furumichi, Kanae Morishima, and Mao Tanabe. 2019. "New Approach for Understanding Genome Variations in KEGG." *Nucleic Acids Research* 47 (D1): D590–95.

Kaufman, R. J., M. V. Davies, L. C. Wasley, and D. Michnick. 1991. "Improved Vectors for Stable Expression of Foreign Genes in Mammalian Cells by Use of the Untranslated Leader Sequence from EMC Virus." *Nucleic Acids Research* 19 (16): 4485–90.

Kildegaard, Helene Faustrup, Deniz Baycin-Hizal, Nathan E. Lewis, and Michael J. Betenbaugh. 2013. "The Emerging CHO Systems Biology Era: Harnessing the 'Omics Revolution for Biotechnology." *Current Opinion in Biotechnology* 24 (6): 1102–7.

Kim, Jee Yon, Yeon-Gu Kim, and Gyun Min Lee. 2012. "CHO Cells in Biotechnology for Production of Recombinant Proteins: Current State and Further Potential." *Applied Microbiology and Biotechnology* 93 (3): 917–30.

Kim, S. J., N. S. Kim, C. J. Ryu, H. J. Hong, and G. M. Lee. 1998. "Characterization of Chimeric Antibody Producing CHO Cells in the Course of Dihydrofolate Reductase-Mediated Gene Amplification and Their Stability in the Absence of Selective Pressure." *Biotechnology and Bioengineering* 58 (1): 73–84.

Klanert, Gerald, Vaibhav Jadhav, Konstantina Chanoumidou, Johannes Grillari, Nicole Borth, and Matthias Hackl. 2014. "Endogenous microRNA Clusters Outperform Chimeric Sequence Clusters in Chinese Hamster Ovary Cells." *Biotechnology Journal* 9 (4): 538–44.

Lam, Jenny K. W., Michael Y. T. Chow, Yu Zhang, and Susan W. S. Leung. 2015. "siRNA Versus miRNA as Therapeutics for Gene Silencing." *Molecular Therapy. Nucleic Acids* 4 (September): e252.

Law, Charity W., Monther Alhamdoosh, Shian Su, Gordon K. Smyth, and Matthew E. Ritchie. 2016. "RNA-Seq Analysis Is Easy as 1-2-3 with Limma, Glimma and edgeR." *F1000Research* 5: 1408.

Law, Charity W., Yunshun Chen, Wei Shi, and Gordon K. Smyth. 2014. "Voom: Precision Weights Unlock Linear Model Analysis Tools for RNA-Seq Read Counts." *Genome Biology* 15 (2): R29.

Lee, Jae Seong, Thomas Beuchert Kallehauge, Lasse Ebdrup Pedersen, and Helene Faustrup Kildegaard. 2015. "Site-Specific Integration in CHO Cells Mediated by CRISPR/Cas9 and Homology-Directed DNA Repair Pathway." *Scientific Reports* 5 (1). https://doi.org/10.1038/srep08572.

Lee, Kyoung Hwa, Soon Young Park, Su Jin Jeong, Da Hyun Jung, Jie Hyun Kim, Seok Hoon Jeong, Il Mo Kang, and Young Goo Song. 2019. "Can Aminoglycosides Be Used as a New Treatment for ? Activity of Recently Isolated." *Infection & Chemotherapy* 51 (1): 10–20.

Lewis, Nathan E., Xin Liu, Yuxiang Li, Harish Nagarajan, George Yerganian, Edward O'Brien, Aarash Bordbar, et al. 2013. "Genomic Landscapes of Chinese Hamster Ovary Cell Lines as Revealed by the Cricetulus Griseus Draft Genome." *Nature Biotechnology* 31 (8): 759–65.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and 1000 Genome Project Data Processing Subgroup. 2009. "The Sequence Alignment/Map Format and SAMtools." *Bioinformatics* 25 (16): 2078–79.

Mack, George S. 2007. "MicroRNA Gets down to Business." *Nature Biotechnology* 25 (6): 631–38.

Masliah-Planchon, Julien, Simon Garinet, and Eric Pasmant. 2016. "RAS-MAPK Pathway Epigenetic Activation in Cancer: miRNAs in Action." *Oncotarget* 7 (25): 38892–907.

Maszczak-Seneczko, Dorota, Teresa Olczak, Piotr Jakimowicz, and Mariusz Olczak. 2011. "Overexpression of UDP-GlcNAc Transporter Partially Corrects Galactosylation Defect Caused by UDP-Gal Transporter Mutation." *FEBS Letters* 585 (19): 3090–94.

McCarthy, Davis J., Yunshun Chen, and Gordon K. Smyth. 2012. "Differential Expression Analysis of Multifactor RNA-Seq Experiments with Respect to Biological Variation." *Nucleic Acids Research* 40 (10): 4288–97.

North, Simon J., Hung-Hsiang Huang, Subha Sundaram, Jihye Jang-Lee, A. Tony Etienne, Alana Trollope, Sara Chalabi, Anne Dell, Pamela Stanley, and Stuart M. Haslam. 2010. "Glycomics Profiling of Chinese Hamster Ovary Cell Glycosylation Mutants Reveals N-Glycans of a Novel Size and Complexity." *The Journal of Biological Chemistry* 285 (8): 5759–75.

Pasquinelli, A. E., B. J. Reinhart, F. Slack, M. Q. Martindale, M. I. Kuroda, B. Maller, D. C. Hayward, et al. 2000. "Conservation of the Sequence and Temporal Expression of Let-7 Heterochronic Regulatory RNA." *Nature* 408 (6808): 86–89.

R Development Core Team. 2003. *The R Reference Manual: Base Package*. Network Theory.

Robinson, Mark D., and Alicia Oshlack. 2010. "A Scaling Normalization Method for Differential Expression Analysis of RNA-Seq Data." *Genome Biology* 11 (3): R25.

Ronda, Carlotta, Lasse Ebdrup Pedersen, Henning Gram Hansen, Thomas Beuchert Kallehauge, Michael J. Betenbaugh, Alex Toftgaard Nielsen, and Helene Faustrup Kildegaard. 2014. "Accelerating Genome Editing in CHO Cells Using CRISPR Cas9 and CRISPy, a Web-Based Target Finding Tool." *Biotechnology and Bioengineering* 111 (8): 1604–16.

Sanchez, Noelia, Paul Kelly, Clair Gallagher, Nga T. Lao, Colin Clarke, Martin Clynes, and Niall Barron. 2014. "CHO Cell Culture Longevity and Recombinant Protein Yield Are Enhanced by Depletion of miR-7 Activity via Sponge Decoy Vectors." *Biotechnology Journal* 9 (3): 396–404.

Sealover, Natalie R., Angela M. Davis, Jeanne K. Brooks, Henry J. George, Kevin J. Kayser, and Nan Lin. 2013. "Engineering Chinese Hamster Ovary (CHO) Cells for Producing Recombinant Proteins with Simple Glycoforms by Zinc-Finger Nuclease (ZFN)-Mediated Gene Knockout of Mannosyl (alpha-1,3-)-Glycoprotein Beta-1,2-N-Acetylglucosaminyltransferase (Mgat1)." *Journal of Biotechnology* 167 (1): 24–32.

Smyth, G. K. n.d. "Limma: Linear Models for Microarray Data." In *Statistics for Biology and Health*, 397–420.

Strotbek, Michaela, Lore Florin, Jennifer Koenitzer, Anne Tolstrup, Hitto Kaufmann, Angelika Hausser, and Monilola A. Olayioye. 2013. "Stable microRNA Expression Enhances Therapeutic Antibody Productivity of Chinese Hamster Ovary Cells." *Metabolic Engineering* 20 (November): 157–66.

Su, Shian, Charity W. Law, Casey Ah-Cann, Marie-Liesse Asselin-Labat, Marnie E. Blewitt, and Matthew E. Ritchie. 2017. "Glimma: Interactive Graphics for Gene Expression Analysis." *Bioinformatics* 33 (13): 2050–52.

Tsuda, E., G. Kawanishi, M. Ueda, S. Masuda, and R. Sasaki. 1990. "The Role of Carbohydrate in Recombinant Human Erythropoietin." *European Journal of Biochemistry / FEBS* 188 (2): 405–11.

Walsh, Gary. 2014. "Biopharmaceutical Benchmarks 2014." *Nature Biotechnology* 32 (10): 992–1000.

Werner, Rolf G., Kristina Kopp, and Michael Schlueter. 2007. "Glycosylation of Therapeutic Proteins in Different Production Systems." *Acta Paediatrica* 96: 17–22.

Wong, Danny Chee Furng, Niki Soo Ching Wong, John Soo Yang Goh, Lee May May, and Miranda Gek Sim Yap. 2010. "Profiling of N-Glycosylation Gene Expression in CHO Cell Fed-Batch Cultures." *Biotechnology and Bioengineering* 107 (3): 516–28.

Yang, Jane, Ronya Primack, Mike Frohn, Wei Wang, Peng Luan, Marc W. Retter, and Gregory C. Flynn. 2015. "Impact of Glycation on Antibody Clearance." *The AAPS Journal* 17 (1): 237–44.

Yang, Zhang, Shengjun Wang, Adnan Halim, Morten Alder Schulz, Morten Frodin, Shamim H. Rahman, Malene B. Vester-Christensen, et al. 2015. "Engineered CHO Cells for Production of Diverse, Homogeneous Glycoproteins." *Nature Biotechnology* 33 (8): 842–44.

Zhang, Peiqing, Susanto Woen, Tianhua Wang, Brian Liau, Sophie Zhao, Chen Chen, Yuansheng Yang, et al. 2016. "Challenges of Glycosylation Analysis and Control: An Integrated Approach to Producing Optimal and Consistent Therapeutic Drugs." *Drug Discovery Today* 21 (5): 740–65.