University of Natural Resources and Life Sciences, Vienna

Department of Biotechnology

Institute of Applied Microbiology

Genome-wide analysis of recombinant protein expression in Chinese hamster ovary cells

A thesis submitted in partial fulfillment of the requirement for the degree of Doctor of Philosophy at the University of Natural Resources and Life Sciences, Vienna.

by

Dipl.-Ing. **Andreas Michael Maccani**

Vienna, July 2014

Supervisors:

Ao.Univ.Prof. Dipl.-Ing. Dr.nat.techn. **Reingard Grabherr**

Dipl.-Ing. Dr.nat.techn. **Wolfgang Ernst**

Univ.Prof. Dipl.-Ing. Dr.nat.techn. **Renate Kunert**

“Scientists are people of very dissimilar temperaments doing different things in very different ways. Among scientists are collectors, classifiers and compulsive tidiers-up; many are detectives by temperament and many are explorers; some are artists and others artisans. There are poet-scientists and philosopher-scientists and even a few mystics.”

Peter B. Medawar, The Art of the Soluble, London: Methuen, 1967, p. 132

Acknowledgements

Writing this PhD thesis would not have been possible without the support of many people.

First of all, I would like to express my sincere gratitude to Professor Reingard Grabherr and Dr. Wolfgang Ernst for giving me the opportunity to conduct my study in their lab and for supervising my work. Professor Grabherr has always had an open door for my concerns and her advice and encouragements were very helpful. Dr. Ernst's scientific support and suggestions were of great value for this thesis. I very much appreciate the collegial collaboration with both of them.

In addition, I am very grateful to Professor Renate Kunert, who guided me through the tedious process of mammalian cell line development, and Dr. Christian Leitner, who provided the knowledge about bioreactor cultivation and supported me during this part of my work.

I am especially grateful to Professor Nicole Borth and Dr. Matthias Hackl for their scientific support and technical assistant during the microRNA analysis.

I am also indebted to all members of the ACIB project 3.5 for sharing their knowledge and experience, especially Professor Diethard Mattanovich for giving me the opportunity to work in this project, Dr. Alexandra Graf and Nadine Tatto for providing bioinformatics support and Nils Landes, my yeast counterpart.

Furthermore, I owe a huge dept of gratitude to all my other colleagues of the Baculo and Kunert group for discussion, encouragement and assistance. They created a positive atmosphere which made it a pleasure to work in their labs. My special thanks to Martina Hofbauer and Willibald Steinfellner for their support in cell culture as well as Stefan Bauernfried for his assistance with the fermentations.

I would like to express my gratitude also to Professor Mark Smales form the University of Kent, Canterbury, UK for being my external supervisor and hosting me in his lab during my research visit. Many thanks also to all members of the Smales lab, who made my stay in Canterbury very enjoyable.

I gratefully acknowledge the Austrian Centre of Industrial Biotechnology (ACIB) for funding my PhD work, as well as the PhD program "BioToP – Biomolecular Technology of Proteins" for additional support.

Lastly, and with no less gratitude, I would like to thank my partner Magdalena and my parents for their patience and enduring support.

Abstract

Chinese hamster ovary (CHO) cells are currently the workhorses of the biopharmaceutical industry for the production of recombinant therapeutic proteins, mainly because of their ability to grow in suspension and to secrete complex proteins with human-like glycosylation. However, although CHO cells have been used for more the 25 years, the cellular processes that control and limit recombinant protein synthesis and secretion are still poorly understood. In addition, other expression systems such as the yeast *Pichia pastoris* are about to enter this field. To compare their capability for recombinant protein secretion, high producing CHO cell lines and *P. pastoris* strains expressing two proteins of different complexity were established and cultivated in comparable bioprocesses. Our results indicate that the protein secretion machinery of CHO cells is much more efficient, and especially the secretion of the more complex protein was strongly impaired in *P. pastoris*.

To investigate recombinant protein production in CHO cells at a physiological level and to provide transcriptomics data for a cross-species comparison with *P. pastoris*, low and high producing CHO cell lines as well as a non-producer were cultivated under steady-state conditions and global gene expression was profiled using a novel CHO-specific microarray. In addition, the expression of microRNAs was analyzed. MicroRNAs are short (~22 nucleotides) non-coding RNAs that play an important role in the regulation of gene expression in higher eukaryotes and therefore may have great potential as cell engineering targets. In the present study, we identified several genes and microRNAs that might be involved in recombinant protein production in CHO cells and consequently may constitute potential targets for cell engineering to improve bioprocess-relevant properties such as productivity or longevity. However, our results also suggest that the reaction of CHO cells to heterologous protein expression is strongly product and/or clone-specific.

Kurzfassung

Chinese hamster ovary (CHO) Zellen sind derzeit die Arbeitstiere der biopharmazeutischen Industrie für die Produktion von rekombinanten therapeutischen Proteinen, hauptsächlich, weil sie dazu befähigt sind, in Suspension zu wachsen und komplexe Proteine mit human-ähnlicher Glykosylierung zu sekretieren. Obwohl CHO Zellen schon seit über 25 Jahren dafür verwendet werden, sind die zellulären Prozesse, welche die Synthese und Sekretion von rekombinanten Proteinen steuern und limitieren, noch sehr schlecht verstanden. Zudem stellen andere Expressionssysteme wie die Hefe *Pichia pastoris* eine interessante Alternative dar. Um die Fähigkeit von CHO Zellen und *P. pastoris* rekombinante Proteine zu sekretieren zu vergleichen, wurden Zelllinien bzw. Stämme entwickelt, die zwei unterschiedlich komplexe Proteine produzieren und diese in vergleichbaren Bioprozessen kultiviert. Unsere Ergebnisse zeigen, dass die Sekretionsmaschinerie von CHO Zellen wesentlich effizienter ist. Besonders die Sekretion des komplexeren Proteins war in *P. pastoris* stark beeinträchtigt.

Um die rekombinante Proteinproduktion in CHO Zellen auf einer physiologischen Ebene zu untersuchen und um Genexpressionsdaten für einen speziesübergreifenden Vergleich mit *P. pastoris* zu generieren, wurden niedrig- und hochproduzierende CHO Zelllinien sowie ein Nichtproduzent unter stationären Bedingungen kultiviert und die Gesamtgenexpression mittels eines neuen CHO-spezifischen Microarrays gemessen. Zusätzlich wurde auch die Expression von microRNAs analysiert. MicroRNAs sind kurze (~22 Nukleotide) nicht-kodierende RNAs, die eine wichtige Rolle bei der Regelung der Genexpression in höheren Eukaryoten spielen. Daher wird ihnen ein großes Potential als Ziel für die gentechnische Optimierung (Cell Engineering) von CHO Zellen zugeschrieben. In der vorliegenden Arbeit wurden etliche Gene und microRNAs identifiziert, die möglicherweise an der rekombinanten Proteinproduktion in CHO Zellen beteiligt sind und daher potenzielle Ziele für Cell Engineering zur Verbesserung bioprozessrelevanter Eigenschaften, wie Produktivität und Langlebigkeit darstellen. Allerdings deuten unsere Ergebnisse auch darauf hin, dass die Reaktion von CHO Zellen auf die Expression von Fremdproteinen stark produkt- und/oder klon-spezifisch ist.

Table of contents

1. Introduction 8
1.1 Expression systems for therapeutic protein production 8
1.2 Chinese hamster ovary (CHO) cells – A historic perspective 11
1.3 Recombinant protein production in CHO cells 13
1.4 Cell engineering of CHO cells 18
1.5 Genomic characteristics of CHO cells and the Chinese hamster 22
1.6 Systems biology approaches towards CHO cell engineering 24
1.6.1 Transcriptomics 24
1.6.2 Proteomics 28
1.6.3 Metabolomics 29
2. Objectives 31
3. Relevant publications 33
3.1 Publication A: Nuclear DNA content of CHO cells 33
3.2 Publication B: Bioprocess comparison – CHO cells vs. *Pichia pastoris* 37
3.3 Publication C: MicroRNA profiling of recombinant CHO cells 57
4. Results and discussion 88
4.1 CHO cells versus *Pichia pastoris*: A system comparison 88
4.1.1 Cell growth and cell density 89
4.1.2 Productivity 89
4.1.3 Product quality 90
4.1.4 Cross-species comparison to improve recombinant protein production 91
4.2 Global mRNA expression profiling in recombinant CHO cell lines 93
4.2.1 Comparable conditions by steady-state cultivation 93
4.2.2 A novel CHO-specific microarray 94
4.2.3 Product/clone dependency of mRNA expression 101
4.2.4 Endoplasmic reticulum (ER) stress and unfolded protein response (UPR) 105

4.2.5 Engineering transcription efficiency 110
4.3 Global miRNA expression profiling in recombinant CHO cell lines 113
4.3.1 Product/clone dependency of miRNA expression 113
4.3.2 Impact of confirmed miRNAs 116
4.3.3 Functional characterization of miRNAs 119
5. Conclusion and perspectives 120
6. References 124
Appendix 142
Supporting information: Validation of CHO-specific microarray 142
List of figures 145
Abbreviations 146
Curriculum vitae 149

1. Introduction

1.1 Expression systems for therapeutic protein production

Human protein therapeutics for the treatment of various previously incurable diseases have revolutionized modern medicine over the last decades. As the isolation from natural sources is often limited, the introduction of recombinant DNA (rDNA) technologies in the mid-1970s represented a milestone for the development of biopharmaceutical drugs. From now on, peptides and proteins could be produced in foreign hosts. In 1982, human insulin produced in *Escherichia coli* (Humulin by Eli Lilly) was the first recombinant drug that gained FDA approval. Since then, many more biopharmaceutical products including hormones, interferons, vaccines as well as monoclonal antibodies have become available. Today, therapeutic proteins are a fast growing multibillion-dollar market. In 2012, the global spending on biological drugs was 169 billion US$ and is expected to reach more than 220 billion US$ in 2017 (Rickwood et al. 2013). In the US, monoclonal antibodies (mAbs) were the highest selling class of biological drugs, followed by hormones and growth factors (Figure 1).

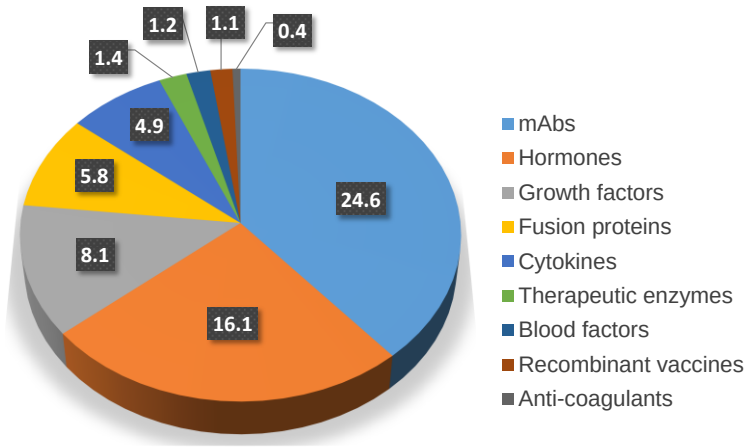


Figure 1. US sales (billion US$) of the top nine categories of biological drugs in 2012. Adapted from Aggarwal (2014)

Today, various production systems are in use for the expression of heterologous proteins ranging from bacterial hosts to transgenic animals. Until January 2009, 151 recombinant pharmaceuticals were approved by the FDA and EMEA. 45 of them were produced in *Escherichia coli*, 28 in *Saccharomyces cerevisiae*, 17 in hybridoma cells, one in transgenic goat, one in insect cells and 59 in mammalian cells (Ferrer-Miralles et al. 2009). Of the 58

approved biopharmaceuticals between 2006 and June 2010, 32 are produced in mammalian cells, 17 in *E. coli* and five in yeast (Walsh 2010).

Therapeutic proteins may require correct post-translational modifications (PTMs) such as glycosylation, disulfide bond formation, phosphorylation and proteolytic processing for biological efficacy and therefore the choice of the most suitable expression system highly depends on the complexity of the product as well as on its ability to conduct PTMs.

Among bacteria, *E. coli* is the predominant host for the production of recombinant proteins. This is mainly due to the vast number and quality of available information because of its role as model organism as well as the availability of well-established molecular biological methods and sophisticated fermentation technologies. This extensive knowledge led to processes which enable very high yields of recombinant proteins up to 50% of total biomass (Panda 2003). Additionally, *E coli* can grow very fast in inexpensive media and fermentation protocols for high cell density cultivation (HCDC) have been established which enabled a biomass concentration of more than 100 g L^{-1} (Lee 1996). However, bacterial hosts lack the ability for correct post-translational processing and in many cases the heterologous proteins aggregate and form inclusion bodies within the cell. Thus, proteins might lose their function unless the correct structure can be restored by *in vitro* refolding. Although genetic engineering efforts made glycosylation also possible in prokaryotic hosts, the authenticity of the glycans is far from eukaryotic systems (Pandhal and Wright 2010). Another major drawback of *E. coli* was the inability to secret proteins into the culture medium. However, genetic engineering also enabled extracellular protein production and yields of more than 2 g L^{-1} have already been reported (Ni and Chen 2009). Additionally, *E. coli* are Gram-negative bacteria and so they contain lipopolysaccharides (LPS) in the outer membrane. These structures are endotoxins which are highly pyrogenic in humans. As even small amounts of endotoxin can cause illness in humans, *E. coli* derived therapeutics require extensive purification to ensure endotoxin-free products.

Yeast expression systems have many of the same advantages as *E. coli*. Molecular biological methods for genetic manipulation are well established and complete genome sequences are available for many strains. Furthermore, they can grow in inexpensive, chemically defined media to very high cell densities. Until 2009, all approved yeast-derived biological drugs were produced in *S. cerevisiae*. But also expression systems based on non-conventional yeasts have been established during the last two decades and especially the methylotrophic yeasts *Hansenula polymorpha* and *Pichia pastoris* have a great potential for the production of recombinant proteins. In 2009, Ecallantide (tradename Kalbitor) was the first therapeutic

derived from *P. pastoris* that was approved by the FDA (Waegeman and Soetaert 2011). The main advantages of yeasts in comparison with *E. coli* are that they can secrete recombinant proteins into the culture broth and that they do not produce endotoxins. In addition, yeasts do not contain oncogenic or viral DNA (Mattanovich et al. 2012). As yeasts are eukaryotes, they have the capacity for protein processing such as protein folding and PTMs. However, the N-linked glycosylation patterns differ significantly from human and consequently can impair serum half-life and immunogenicity of therapeutic proteins (Dasgupta et al. 2007; Sola and Griebenow 2009). The oligosaccharide chains added to the proteins mainly consist of mannose units. In addition, *S. cerevisiae* often hyperglycosylates the proteins by adding up to 150 mannose residues per glycosylation site, whereas the majority of oligosaccharides added in *P. pastoris* contain only 8–16 mannose units (Grinna and Tschopp 1989; **Maccani et al. 2014**). The absence of human typical sugar residues such as sialic acid and the high mannose content are the main drawbacks of yeasts compared to mammalian cells. Hence, a lot of effort has been put into the engineering of strains that are able to synthetize human-like N-glycans in *P. pastoris* (Hamilton et al. 2006; Potgieter et al. 2009; Ye et al. 2011). However, no glycosylated therapeutic proteins produced in yeast have entered the market to date.

For this reason, mammalian cells are the expression systems of choice today, because they are able to produce complex therapeutic proteins such as mAbs which require proper folding, human-like PTMs or multimeric assembly. A variety of mammalian cell lines have been established including Chinese hamster ovary (CHO), mouse myeloma (NS0 and SP2/0), baby hamster kidney (BHK), human embryonic kidney (HEK-293) and human embryonic retinoblast (PER.C6) cells. Most of these cell lines were adapted to grow in suspension which makes them highly suitable for large scale bioreactor cultivation (up to 20,000 L) commonly reaching yields of 1–8 g L^{-1} in industrial processes (Beck and Reichert 2012). Monoclonal antibodies are the largest and fastest growing class of therapeutic proteins with several hundred in different stages of clinical trials. Of the 28 mAbs on the European Union or US market in 2012, 43% were produced in CHO cells, 25% in SP2/0 cells, 18% in NS0 cells and 7% in hybridomas (Reichert 2012), which clearly shows that CHO cells are currently the workhorses of the biopharmaceutical industry. This will most likely not change in the near future as the cell line is well characterized and has been used very successfully for a wide range of products over the last three decades. Hence, it is much easier and very likely more rapid to gain regulatory approval of new biological drugs derived from CHO cells than from a less well-known or an entirely new cell line.

1.2 Chinese hamster ovary (CHO) cells – A historic perspective

Chinese hamsters (*Cricetulus griseus*) belong to the rodent family of *Cricetidae* and are native to the desert of northern China and Mongolia (Figure 2A). Almost 100 years ago in 1919, Chinese hamsters were already used in laboratory for typing pneumococci. In the early 1920's, they became a valuable tool for the study of leishmaniasis (also known as kala-azar or black fever) as they were known as carriers of *Leishmania*, a parasite causing this deadly disease. In 1948, Dr. C. H. Hu of the Peking Union Medical College gave some specimens to Dr. Robert Watson of the Harvard Medical School who smuggled them into the United States. The Chinese government thought that the hamsters were part of a biological weapons program and so both scientists were accused of war crimes, which led to the imprisonment of Dr. Hu. (Jayapal et al. 2007; Yerganian 1985). In the 1950s, it was noticed that Chinese hamsters have fewer and therefore larger chromosomes (2n = 22) than mice (2n = 40) and rats (2n = 42), which made them an excellent model for radiation cytogenetic and tissue culture studies (Tonomura and Yerganian 1956; Yerganian 1952). The birth of the original CHO cell line was in 1957 when Dr. Theodore T. Puck of the University of Colorado Medical Center in Denver isolated an ovary from a female Chinese hamster provided by Dr. George Yerganian of the Boston Children's Cancer Research Foundation and successfully recovered a spontaneously immortalized population of cultured fibroblast cells (Tjio and Puck 1958). Cultured adherent CHO cells are illustrated in Figure 2B. As they were easy to maintain, CHO cells were thereafter used in numerous medical and cell biology studies, making them the mammalian counterpart of *E. coli* as a laboratory model. In the 1960, Puck and others used mutagenesis to generate auxotrophic cells that required certain nutrients for maintaining growth and viability (Kao and Puck 1967; Puck and Kao 1967). The use of CHO cells in the biopharmaceutical industry began in the early 1980s after Dr. Gail Urlaub and Dr. Lawrence A. Chasin of the Columbia University in New York isolated cell lines (DUKX-B11 and DG44) lacking the metabolic enzyme dihydrofolate reductase (DHFR) after mutagenesis and selection of thymidine auxotrophic cells (Urlaub and Chasin 1980; Urlaub et al. 1983). This genetic defect was then utilized to generate stable production cell lines by introducing the transgene together with a functional copy of the DHFR gene (*dhfr*) as selection marker. In addition, this method enabled the amplification of the transgene copy number to increase the productivity using methotrexate (MTX) which is an antagonist of DHFR. In the mid-1980s, recombinant CHO cell lines were adapted to grow in suspension in serum-free media (Figure 2C), which finally cleared the way for the industrial application (Wurm and Hacker

2011). In 1987, the tissue plasminogen activator for the treatment of myocardial infarctions was the first CHO-expressed therapeutic protein that gained approval for clinical application. Since then dozens biologic drugs produced in CHO cells including many monoclonal antibodies have reached the market. The most recent milestones were the publications of the first CHO genomic sequences in 2011 (Hammond et al. 2012a; Xu et al. 2011b) and Chinese hamster genome sequences in 2013 (Brinkrolf et al. 2013; Lewis et al. 2013), which heralded a new era of CHO research.

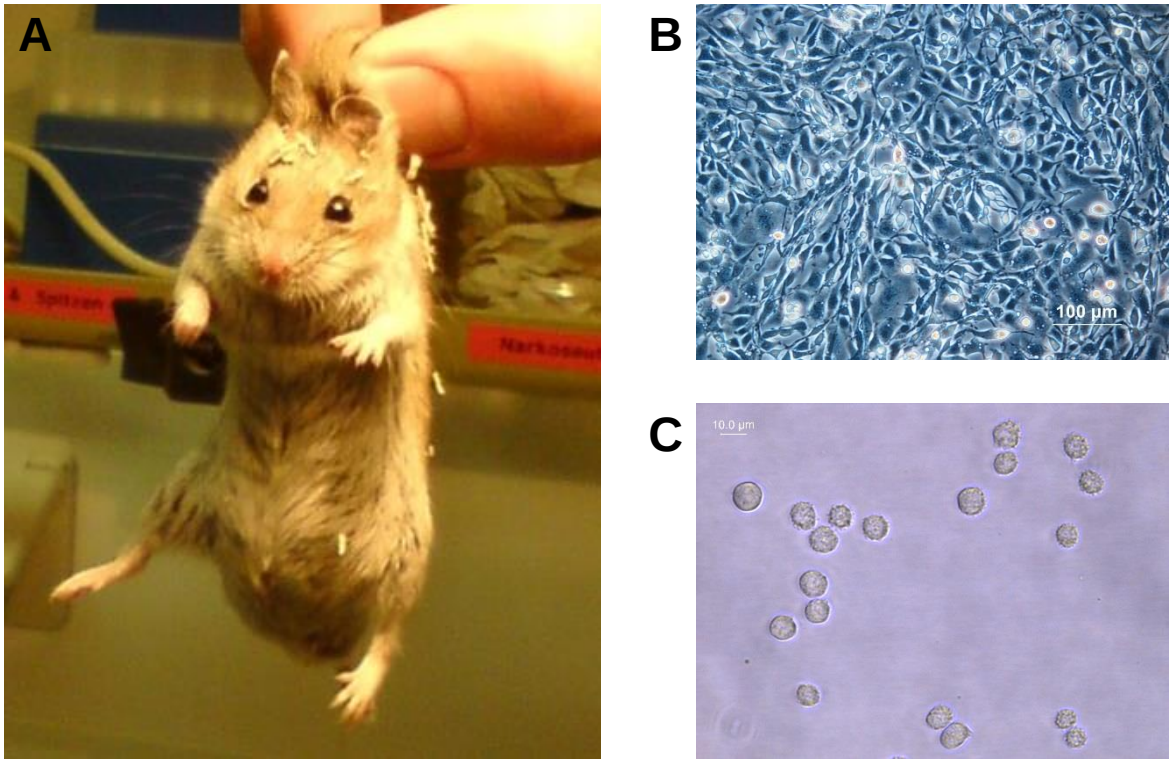


Figure 2. From Chinese hamster to Chinese hamster ovary (CHO) suspension cells. Chinese hamsters (*Cricetulus griseus*) belong to the rodent family of *Cricetidae* and are native to the desert of northern China and Mongolia (**A**). In 1957, Dr. Theodore T. Puck isolated the original CHO cell line from the ovaries of a female Chinese hamster and successfully recovered a spontaneously immortalized population of cultured fibroblast cells. The cells were adherent requiring a surface to grow (**B**). In the mid-1980s, recombinant CHO cell lines were adapted to grow in suspension in serum-free media (**C**), which finally cleared the way for the industrial application.

1.3 Recombinant protein production in CHO cells

Chinese hamster ovary (CHO) cells are the predominant mammalian hosts for the generation of stable high producing clones in the biopharmaceutical industry for various reasons. They can easily be adapted to grow in suspension which enables large scale cultivation. Furthermore, several systems for gene amplification to generate clones with high specific productivity are well established. CHO cells can grow in serum-free, protein-free and chemically defined media which is important to ensure reproducibility between different batches. In addition, this reduces the risk of contaminations with viruses or prions which might be introduced by bovine or human serum, and it also simplifies downstream processing because of the lower content of protein contaminations. An advantage over human cell lines is that CHO cells are safer in regard to human viruses. In a study from 1989, 44 human pathogens including influenza, polio, herpes, measles and human immunodeficiency virus (HIV) were tested and the majority of them did not replicate in CHO cells (Wiebe et al. 1989). Beside safe products, CHO cells are also able to produce complex therapeutic proteins that are biological active and stable in humans, which require correct protein folding and post-translational modifications (PTMs) such as human-like glycosylation. The glycoforms on the Fc domain of IgGs produced in CHO cells have the complex-type biantennary structures which are typically found on human IgGs (Beck and Reichert 2012; **Maccani et al. 2014)**. In addition, the levels of immunogenic glycoforms with terminal galactose-α-1,3-galactose or N-glycolylneuraminic acid, which may appear as traces, are lower compared to those frequently found on glycoproteins produced in the murine cell lines NS0 or Sp2/0 (Beck and Reichert 2012; Bosques et al. 2010).

A major drawback of CHO cells (and immortalized cell lines in general) is the high degree of genetic and phenotypic diversity between different cell lines, which is a result of the long and uncontrolled history of these cells including changes from chemical and radiation mutagenesis (e.g. DUKX-B11 and DG44) together with the cultivation under highly diverse conditions in the different labs (Wurm 2013).

In the mid-1980s, when industry started to use CHO cells for recombinant protein production, they were regarded as a not very effective expression system yielding only 50–100 mg L^{-1} in typical batch processes. However, due to the establishment of gene amplification systems as well as vector, media and process optimization, antibody concentrations of more than 1 g L^{-1} are typically reached in industrial fed batch processes today (De Jesus and Wurm 2011).

The currently applied platforms for CHO cell line development are mainly based on random integration and gene amplification using either the dihydrofolate reductase (DHFR) or the glutamine synthetase (GS) system. The DHFR system originates from the early 1980s (Kaufman and Sharp 1982) and can be applied for the DHFR-deficient cell lines CHO DUKX-B11 and DG44 which are triple auxotrophs for glycine, hypoxanthine and thymidine. The cells are co-transfected with the gene of interest (GOI) and a functional copy of the DHFR gene (*dhfr*) which enables the selection of recombinant clones in media free from hypoxanthine and thymidine. The system makes use of methotrexate (MTX), a folic acid analog that blocks DHFR activity. A stepwise increase of the MTX concentration in the culture medium results in an amplification of the transfected *dhfr* and the GOI. The GS system was initially developed for NS0 cells (Bebbington et al. 1992), but was also adapted for CHO cells and is commonly used for CHO-K1 cells, a subclone of the original CHO cell line. The principal is similar to the DHFR system. Recombinant cells are obtained by co-transfection of a functional GS gene together with the GOI and selection in the absence of glutamine. The application of methionine sulfoximine (MSX), a GS inhibitor, enables the amplification of the GS gene and the GOI. The principal of cell line development using these technologies is illustrated in Figure 3.

Random gene integration and gene amplification results in very heterogeneous cell populations which requires single cell cloning such as limiting dilution followed by the screening of thousands of clones in order to identify high producing ones. The efficiency of the screening process can be improved by using high-throughput techniques such as fluorescence-activated cell sorting (FACS)-based systems (Yoshikawa et al. 2001) or the ClonePix FL system (Dharshanan et al. 2011). However, the development of production cell lines is still a very time-consuming and labor-intensive process which takes at least six month (Lai et al. 2013).

Over the past years, various vector engineering strategies based on site-specific integration and cis-acting elements have been developed that facilitate the establishment of stable high producing cell lines (Kim et al. 2012; Lai et al. 2013). The traditionally used random integration of the transgene into the chromosomes leads to a huge variance of the specific productivity between the individual clones (chromosomal position effect). In addition, it might favor the instability of the specific productivity which is frequently observed during long-term cultivation and is likely caused by the chromosomal environment due to gene silencing via methylation or chromosomal rearrangement (Barnes et al. 2003). But high producing clones can also be generated in a reproducible and predictable manner using site-

specific integration based on recombinase-mediated cassette exchange (RMCE). This technique, however, requires the identification of a chromosomal locus that allows a stable and high expression of the GOI, followed by the establishment of a host cell line containing the recognition sequence of a specific recombinase in the identified locus. The two site-specific recombinases Cre and Flp, which recognize the loxP and FRT sequence respectively, are most commonly used. The first use of targeted gene integration in CHO cells was reported by Kito et al. (2002), who used the Cre/loxP system to generate mAb producing cell lines. Another recombinase-based gene targeting system uses an artificial mammalian chromosome that contains 50–70 recombination acceptor sites together with a mutant lambda integrase, which allows the incorporation of multiple copies of the GOI (Kennard et al. 2009). Another strategy is the use of large vectors such as bacterial artificial chromosomes (BACs) which contain regions of open chromatin that is transcriptionally active. BACs have a large cloning capacity (up to 300kB) and due to their large size, transcription from BACs is less affected by the surrounding chromatin (Blaas et al. 2009; Mader et al. 2013).

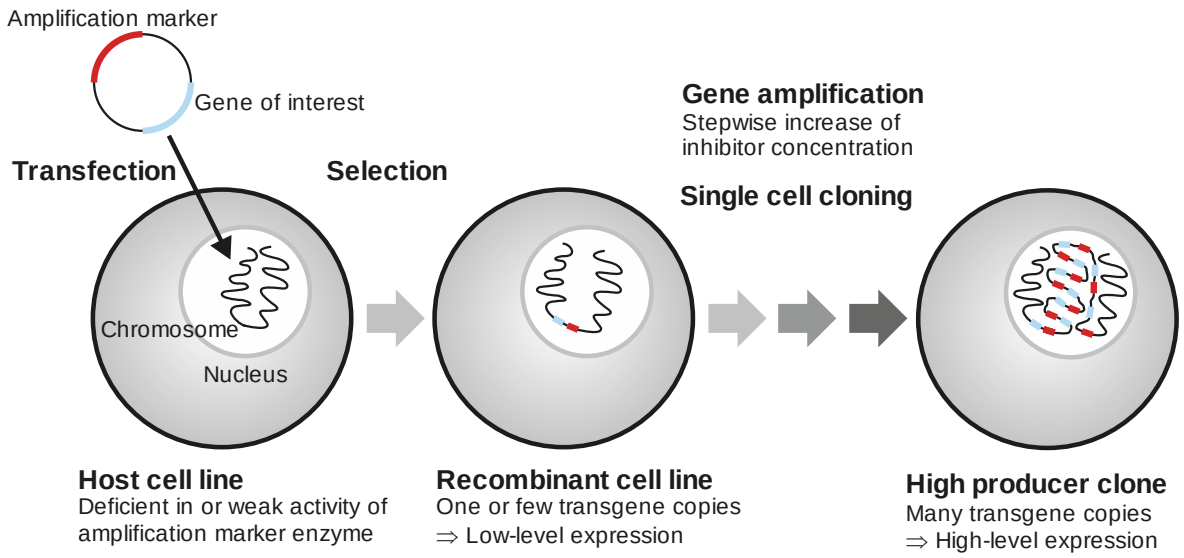


Figure 3. Typical process to develop a recombinant high producer CHO cell line by gene amplification. An auxotroph host cell line (e.g. dihydrofolate reductase deficient CHO cells) is transfected with a vector containing the gene of interest (GOI) and an amplification (selection) marker (e.g. *dhfr*) to obviate the need for the nutrients the host cell is auxotroph for (e.g. glycine, hypoxanthine and thymidine). This allows the selection of recombinant cell lines with one or few copies of the GOI. Applying increasing amounts of an appropriate inhibitor of the amplification marker (e.g. methotrexate (MTX)) leads to the amplification of the amplification marker as well as genetically linked sequences including the GOI. Repeating rounds of gene amplification and single cell cloning (e.g. limiting dilution) enables the establishment of clonal high producer cell lines. Adapted from **Maccani et al. (2013)**

Other approaches use cis-acting elements which can remodel chromatin to keep the GOI in an active conformation. Matrix attachment regions (MARs) and scaffold attachment regions (SARs) maintain a transcriptionally active chromatin by creating chromatin loops and it was demonstrated that MARs and SARs flanking the GOI are able to improve recombinant gene expression in CHO cells (Girod et al. 2005; Kim et al. 2005; Kim et al. 2004; Zahn-Zabal et al. 2001). Ubiquitous chromatin opening elements (UCOEs) are cis-acting elements that maintain chromatin in an open configuration. The incorporation of UCOEs upstream of the promoter enables gene expression that is independent of the integration site and resistant to silencing and it was shown to reduce the screening effort (Benton et al. 2002; Ye et al. 2010).

An alternative to increasing selection pressure by increasing the drug concentration is the attenuation of the selection/amplification marker. Consequently, lower drug concentrations should be sufficient to obtain high producer clones, which might also reduce potential side-effects of the drug such as reduced cell growth. One strategy is the mutation of the selection marker to reduce its activity. Reducing the affinity of the selection marker neomycin phosphotransferase II by introducing a mutation led to an increased productivity of mAbs expressing CHO cells (Sautter and Enenkel 2005). Another strategy is to impair the expression of the selection marker. This can be done by using a weak promoter to drive transcription of the selection marker (Niwa et al. 1991). A different strategy was used by Westwood et al. (2010) who improved product yields of recombinant CHO cells by codon deoptimization of the selection marker *dhfr* in order to lower translational efficiency. Alternatively, mRNA and protein destabilizing elements can be used to reduce the expression of the selection marker. Using short AU-rich elements (AREs) and the mouse ornithine decarboxylase (MODC) PEST region as destabilizing elements resulted in a significant improvement of the specific product secretion rate (q_P) using the DHFR/MTX amplification system (Ng et al. 2007). Internal ribosome entry site (IRES) elements are frequently used to directly link the GOI and the selection marker, which ensures expression of the GOI upon selection (Lai et al. 2013).

The most essential element for gene transcription, however, is the promoter region. High-level protein expression requires strong constitutive promoters. The human and mouse cytomegalovirus (CMV) major immediate-early (MIE) enhancer/promoter is most commonly used. But also the Simian virus 40 (SV40) immediate-early promoter and the Rous sarcoma virus (RSV) long terminal repeat (LTR) promoter are frequently applied to drive heterologous gene expression in mammalian cells (Makrides 1999). But also the application of CHO endogenous transcription regulatory elements has been described. The 5’ and 3’ flanking

regions of the CHO-derived elongation factor-1α (CHEF-1) gene were used for heterologous protein expression in CHO cells as well as various other mammalian cell lines, resulting in significantly higher expression levels compared to the CMV promoter (Running Deer and Allison 2004). More recently, the use of synthetic promoters has been proposed (Brown et al. 2014; Grabherr et al. 2011).

1.4 Cell engineering of CHO cells

A multitude of cell engineering approaches to enhance productivity of recombinant Chinese hamster ovary (CHO) cells have been described. The aim of these strategies was mainly to improve the time integral of viable cell concentration (IVCC) and/or the specific product secretion rate (q_P). The IVCC can be in increased by increasing maximum viable cell density, specific growth rate (μ) and/or culture longevity.

Delaying programmed cell death is a very common cell engineering approach to increase IVCC. During cultivation, cell death is induced by various types of stress including nutrient depletion, accumulation of toxic by-products, increased osmolarity and sheer stress. Preventing apoptosis by overexpression of anti-apoptotic proteins such as Bcl-2, Bcl-x_L and Mcl-1 (Chiang and Sisk 2005; Kim and Lee 2001; Majors et al. 2009; Meents et al. 2002a) or down-regulation of pro-apoptotic proteins such as Bak and Bax (Cost et al. 2010) was shown to improve recombinant protein production. As the apoptotic signal is mediated by a caspase-cascade system, the suppression of caspases using antisense RNA or small interfering RNA (Kim and Lee 2002; Sung et al. 2007) or caspase inhibitors such as X-linked mammalian inhibitor of apoptosis (Sauerwald et al. 2002) was also beneficial to delay apoptosis.

Engineering of CHO cells with proliferation-related proteins can be used to increase μ and/or the maximum viable cell concentration (VCC). Enhanced expression of cyclin-dependent kinase like 3 (CDKL3) resulted in elevated cell proliferation (Jaluria et al. 2007), and the overexpression of the cell cycle transcription factor E2F-1 was shown to increase the maximum VCC (Majors et al. 2008). An increased proliferation rate and maximum VCC was achieved by overexpression of c-Myc, which promotes an earlier transition from G1-S by shortening the G1 phase of the cell cycle (Kuystermans and Al-Rubeai 2009). A very interesting target for cell engineering is the mammalian target of rapamycin (mTOR) signaling pathway which plays important roles in many cellular processes including cell proliferation, apoptosis, metabolism, and protein synthesis. Dreesen and Fussenegger (2011) showed that the ectopic expression of human mTOR can improve cell growth, proliferation, viability, robustness and q_P of recombinant CHO cells.

The major bottlenecks preventing mammalian cells from fully exploiting their physiologic production capacity are most likely translation or post-translational processes such as protein folding (Mohan et al. 2008). Hence, engineering the chaperone network might improve the capacity for recombinant protein production of CHO cells. An increase in q_P was observed in thrombopoietin (TPO) producing CHO cells by overexpression of ERp57, an isoform of

protein disulfide isomerase (Hwang et al. 2003), and co-overexpression of calnexin and calreticulin (Chung et al. 2004). The overexpression of protein disulfide isomerase (PDI) could increase the product secretion rate of antibody producing CHO cell lines (Borth et al. 2005; Mohan et al. 2007), however failed to enhance productivity of CHO cells expressing other proteins such as TPO, interleukin-15 (IL-15) and a tumor necrosis factor receptor:Fc fusion protein (Davis et al. 2000; Mohan et al. 2007). Other strategies target pathways of the unfolded protein response (UPR) which are induced upon endoplasmic reticulum (ER) stress. The ectopic expression of the human transcription factor XBP1s (spliced X-box binding protein 1) increased the secretion capacity of recombinant CHO cells (Tigges and Fussenegger 2006). In another study, the same approach was only successful in a transient transfection system but not when applied for stable cell lines (Ku et al. 2008). The overexpression of two other prominent transcription factors of the UPR, ATF4 and GADD34, improved the q_P of antithrombin III-producing CHO cells (Ohya et al. 2008; Omasa et al. 2008). Low temperature shift (37°C to 30–33°C) is a common cultivation strategy to increase q_P of recombinant CHO cells, which leads to a change in expression of many cold-stress proteins such as the cold-inducible RNA-binding protein (CIRP). The stable overexpression of CIRP was shown to improve the productivity of interferon-γ expressing CHO cells at 37°C (Tan et al. 2008). In a recent study, Le Fourn et al. (2014) found that the improper cleavage of the light chain signal peptide of a "difficult-to-express" immunoglobulin (infliximab) resulted in precipitation of the light chains in an insoluble cellular fraction and consequently to low secretion levels. Overexpression of components of the signal recognition particle (SRP) pathway such as the signal receptor protein SRP14, however, restored proper protein processing and secretion leading to high yields of the "difficult-to-express" protein and also increased the production of an "easy-to-express" antibody (trastuzumab). Another potential bottleneck of in the secretory pathway might be related to vesicle trafficking. Soluble NSF (N-ethylmaleimide-sensitive factor) receptors (SNAREs) and Sec1/Munc18 (SM) proteins play a key role in membrane fusion between different compartments. Ectopic expression of the SNAREs SNAP-23 and VAMP as well as the SM proteins Sly1 and Munc18c was able to increase the secretion capacity of various recombinant CHO cell lines (Peng et al. 2011; Peng and Fussenegger 2009).

The accumulation of toxic byproducts such as ammonium and lactate is known to limit cell growth and VCC (Lao and Toth 1997; Xing et al. 2008). Ammonium is a waste product of the glutamine catabolism and its production could be reduced in glutamine-free cultures using engineered recombinant CHO cells expression glutamine synthetase (Zhang et al. 2006). In

another study, the ammonium production of recombinant CHO cell lines was reduced by the overexpression of carbamoyl phosphate synthetase I (CPS I) and ornithine transcarbamoylase (OTC) which catalyze the first and second steps of the urea cycle (Park et al. 2000). Lactate accumulation is a result of excessive glucose consumption. Repressing lactate dehydrogenase (LDH) and/or glycerol-3-phosphate dehydrogenase (GPDH) decreased the lactate content in CHO cell cultures (Jeong et al. 2006). In another study, the fructose-specific transporter GLUT5 was expressed in CHO cells which enabled the utilization of fructose instead of glucose. The cultivation of these cells in media containing fructose in place of glucose led to reduced sugar consumption and lactate production because the GLUT5 fructose transporter is less specific for its substrate (Wlaschin and Hu 2007).

Engineering approaches to alter cellular pathways commonly apply the stable overexpression or silencing of individual genes. The stable overexpression of one or more genes, however, constitutes an additional burden to the translational machinery of the cell. Hence, microRNAs (miRNAs) have been proposed as potential targets for cell engineering as they are non-coding RNAs that can modify the expression of entire groups of genes (Hackl et al. 2012b; Jadhav et al. 2013; Müller et al. 2008). MicroRNAs are short (~22 nt) endogenous RNAs that play an important role in the regulation of gene expression (Bartel 2004). They are predominantly transcribed by RNA Polymerase II (RNA Pol II), processed by Drosha and Dgcr8, and exported to the cytoplasm where the ~70 nt precursor hairpin (pre-miRNA) is cleaved by Dicer resulting in a miRNA duplex (Figure 4). One strand (mature miRNA) associates with Argonaute (AGO) proteins and forms a miRNA-induced silencing complex (miRISC) which recognizes the target mRNAs predominantly by binding to the 3' untranslated region (UTR) through imperfect base-pairing (Krol et al. 2010). MicroRNAs act in a post-transcriptional manner by decreasing translational efficiency and/or transcript levels. A single miRNA can repress hundreds of different mRNAs and thereby regulates entire gene networks (Hobert 2008). They play crucial roles in a wide range of biological processes including development, proliferation, differentiation, apoptosis and metabolism (Bartel 2004; He and Hannon 2004). In mammalian cells, miRNAs are predicted to regulate or fine-tune gene expression of ~50% of all protein-coding genes (Koh et al. 2009). The stable inhibition of mmu-miR-466h-5p which reduces the expression of several anti-apoptotic genes was shown to improve apoptosis resistance in CHO cells (Druz et al. 2013). Jadhav et al. (2014) observed enhanced growth and an increase of q_P with an engineered recombinant CHO cell line overexpressing miR-17. Another group demonstrated that the overexpression of miR-7 can increase the q_P of recombinant CHO cells (Barron et al. 2011).

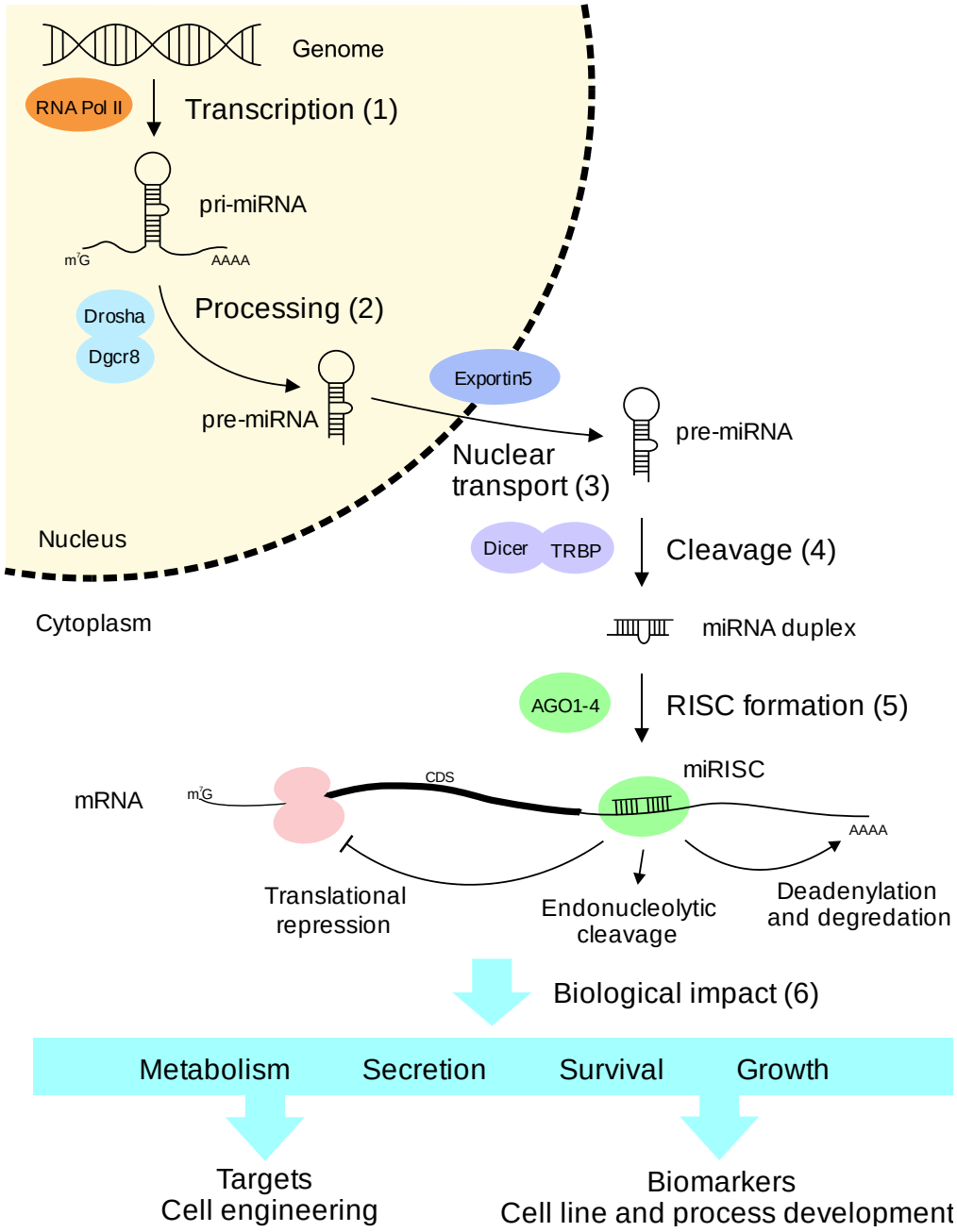


Figure 4. MicroRNA biogenesis and their biotechnological application. MicroRNAs (miRNAs) are small non-coding RNAs that are transcribed predominantly by RNA Polymerase II (RNA Pol II) from intergenic regions or introns (1). The primary miRNA transcripts (pri-miRNAs) are processed by the Drosha/Dgcr8 complex (2), resulting in ~70 nt long precursor miRNAs (pre-miRNAs) that are exported into the cytoplasm via Exportin5 (3). In the cytoplasm, the RNase Dicer in complex with the double-stranded RNA-binding protein TRBP cleaves the pre-miRNA hairpin into the ~22 nt long mature miRNA duplex (4). One strand of the mature miRNA forms together with Argonaute proteins (AGO) the RNA-induced silencing complex (RISC), which is directed to a specific mRNA where miRISC predominantly binds to the 3' untranslated region (3' UTR) by imperfect base-pairing (5), resulting in translational repression and/or mRNA decay. It has been shown that miRNAs play crucial roles in the regulation of a wide range of biological processes (6), which suggests the application of miRNAs as targets for cell engineering as well as biomarkers for cell line and bioprocess development. Illustration based on Krol et al. (2010) and Hackl et al. (2012a)

1.5 Genomic characteristics of CHO cells and the Chinese hamster

In 2011, the publication of a draft genomic sequence of the CHO-K1 ancestral cell line was a major breakthrough for the genomic characterization as well as for cellular and metabolic engineering of Chinese hamster ovary (CHO) cells. The assembly of paired-end sequence reads from an Illumina sequencer comprised 2.45 Gb of the genome with 24,383 predicted genes (Xu et al. 2011b). The authors of this study estimated the CHO-K1 genome size to be 2.6 Gb, which was a very helpful information to improve the estimation of the nuclear DNA content of CHO cells. The unreplicated nuclear DNA content of one cell is required to determine the absolute transgene copy number per cell and was estimated to be 5.4 pg in CHO cells (**Maccani et al. 2013**). Xu et al. (2011b) also found that several key viral entry receptors for viruses are not expressed in the CHO-K1 cell line, which can explain the high resistance to viral infections of CHO cells. In addition they identified homologs for 99% of the human glycosylation-related genes. Although only 53% of them were found to be expressed under exponential growth conditions, it can explain the similarity of glycan structures obtained from human and CHO cells. Furthermore, these results suggest a great potential of CHO cells for glycoengineering (Xu et al. 2011b).

However, the value of genome data from an ancestral cell line might be limited, as a considerable genomic heterogeneity between different CHO cell lines has been observed in various cytogenetic studies. Hence, it is unknown whether genes are still present in the cell lines used today. The diploid genome of the Chinese hamster has 22 chromosomes (10 pairs of autosomes and 2 sex chromosomes), but CHO cells have an altered karyotype. Deaven and Petersen (1973) analyzed the parental cell line of CHO-K1 and found that only 8 of the determined 21 chromosomes appear to be normal compared to the euploid Chinese hamster chromosomes and the remaining 13 chromosomes (Z group) contain translocations, deletions and pericentric inversions. Furthermore, various karyotyping studies have shown that immortalized cell lines undergo continuously genomic modifications, which very likely account for the significant phenotypic differences between different cell lines (Wurm 2013). Chromosomal heterogeneity was already observed in the ancestor of CHO-K1 where a considerable number of the cells had chromosome numbers deviating from 21 (Deaven and Petersen 1973). Derouazi et al. (2006) karyotyped the DG44 host cell line as well as 16 DG44-derived recombinant cell lines. The DG44 host had a total of 20 chromosomes of which seven were normal, five could be assigned to the Z group, seven were derivative chromosomes and two were marker chromosomes. Remarkably, for more than 60% of the

recombinant cell lines, genomic rearrangements relative to DG44 were observed which were specific to each cell line (Derouazi et al. 2006). In a more recent study, Cao et al. (2012) constructed a BAC-based physical map of the CHO DG44 chromosomes and made a karyotypic comparison between DG44, CHO-K1 and primary Chinese hamster cells. Supporting earlier findings, they observed extensive chromosomal rearrangements compared to the diploid hamster genome as well as between DG44 and CHO-K1. Nevertheless, the draft genomic sequence of the CHO-K1 cell line will facilitate the sequencing of other CHO cell lines as well as the establishment of tools (e.g. CHO-specific mRNA microarrays) for a better understanding of CHO cell physiology.

Considering the genetic heterogeneity among CHO cell lines, a well-defined reference genome that comprises all native Chinese hamster genes and regulatory elements would be very beneficial. To address this issue two draft genome sequences of the Chinese hamster were published recently (Brinkrolf et al. 2013; Lewis et al. 2013). Brinkrolf et al. (2013) isolated individual chromosomes of a male Chinese hamster by flow cytometry following separate Illumina sequencing reactions, which allowed the assignment of the sequence assemblies to specific chromosomes. Lewis et al. (2013) assembled a 2.4 Gb genomic sequence of a female Chinese hamster consisting of 24,044 genes. In addition, they sequenced the genomes of six CHO cell lines derived from CHO-K1, DG44 and CHO-S. Comparing the Chinese hamster and CHO-K1 genomes revealed a 99% overlap in gene content. However, numerous genomic variations were detected between different CHO cell lines as well as relative to the Chinese hamster including single-nucleotide polymorphisms (SNPs), short insertions and deletions (indels), gene copy number variations (CNVs) and missing genes (Lewis et al. 2013). Interestingly, they found that several processes associated with proliferation and immortalized phenotypes in cancer cells including the WNT and mTOR signaling pathways and autophagy were insulated from mutations. In addition, they found duplications of many apoptotic genes.

In order to bundle all CHO genomic information and make them publicly accessible, a website (www.chogenome.org) was created (Hammond et al. 2012a). The website hosts the latest updates of CHO genome information and provides genome browsing tools.

1.6 Systems biology approaches towards CHO cell engineering

The various omics-based tools can provide valuable information on a systems biology level to characterize Chinese hamster ovary (CHO) cell factories (Figure 5). These techniques enable the study of the cellular physiology on different levels (transcriptome, proteome and metabolome), and may play an important role in identifying bottlenecks, missing activities and regulatory mechanisms that prevent a desired phenotype such as high productivity, robustness and longevity. In addition, these data will provide the foundation for mathematical models that describe CHO phenotypes as well as for innovative cell and metabolic engineering approaches.

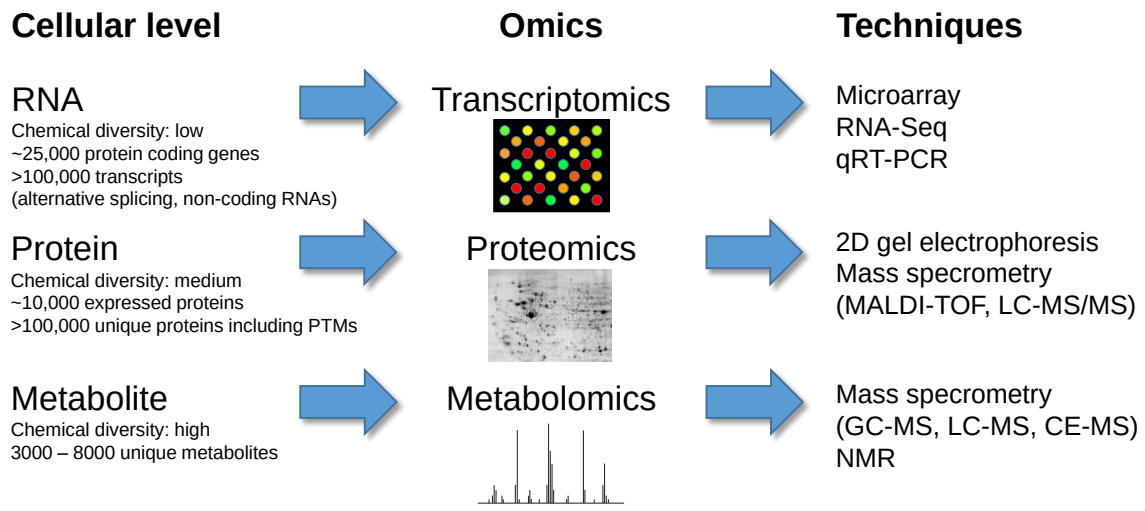


Figure 5. Overview of omics techniques. Transcriptomics aims to measure all transcribed RNAs (transcriptome), proteomics the complete set of expressed proteins (proteome) and metabolomics the complete set of metabolites (metabolome) in a cell. CHO cells have about 25,000 protein coding genes (Xu et al. 2011b), but the number of transcripts is much larger (>100,000) because of splicing variants and non-coding RNAs such as microRNAs. About 10,000 gene products are present in a mammalian cell (Cox and Mann 2007), however, mammalian cells may potentially produce much more unique proteins due to post-translational modifications (PTMs) such as glycosylation and phosphorylation (Jensen 2004). Metabolites are chemically highly diverse. The number of unique metabolites is estimated to be between 3000 and 8000 in mammalian cells (Zhang et al. 2011).

1.6.1 Transcriptomics

The transcriptome refers to all transcribed RNAs (e.g. mRNAs and miRNAs) in a cell under a specific physiological condition. Several methods have been used to study the transcriptome of CHO cells including DNA microarrays, quantitative reverse transcription polymerase chain reaction (qRT-PCR) and more recently transcriptome sequencing (RNA-Seq). DNA microarray analysis is the most comprehensive and powerful technique used to assess

transcriptional changes and differences. The suitability of this method, however, relies on the available genomic sequence information. Until the first genomic sequences of CHO cells became publicly available in 2011 (Xu et al. 2011b), analyses were based on incomplete CHO-specific microarrays constructed from expressed sequence tag (EST) data (Kantardjieff et al. 2009; Melville et al. 2011; Nissom et al. 2006; Wlaschin et al. 2005) which were, however, not publicly available, and mouse or rat-derived DNA microarrays (Baik et al. 2006; Ernst et al. 2006; Yee et al. 2008b). Comparative DNA microarray analyses were done to study the effects of specific product secretion rate (q_P)-enhancing culture conditions such as low temperature (Baik et al. 2006; Yee et al. 2009), sodium butyrate treatment (De Leon Gatti et al. 2007; Yee et al. 2008a) and high osmolarity (Shen et al. 2010). In a study combining low culture temperature (33°C) and sodium butyrate treatment, Kantardjieff et al. (2010) observed that elements of the secretory pathway, including Golgi apparatus, cytoskeleton protein binding and small GTPase-mediated signal transduction were enriched, suggesting that the increased productivity under treatment is the result of an elevated cellular secretory capacity. Doolan et al. (2010) compared fast and slow growing CHO cells in order to identify genes associated with a high specific growth rate (μ) phenotype. They identified several candidates including the valosin-containing protein (VCP) that were involved in the regulation of cell growth. Wong et al. (2006) used DNA microarrays to profile gene expression during batch and fed-batch CHO cell cultures, in order to investigate induction and regulation of apoptosis at the later stage of cultivation. They found that apoptosis was predominantly regulated via the death receptor- and mitochondria-mediated signaling pathways rather than by the endoplasmic reticulum-mediated apoptosis signaling pathway. Only few microarray studies aimed at identifying genes associated with a high q_P phenotype (Kang et al. 2014; Nissom et al. 2006; Trummer et al. 2008; Vishwanathan et al. 2014).

As the cost for next generation sequencing (NGS) continuously decreases, it is anticipated that RNA-Seq will become an interesting alternative to microarrays for the generation of comprehensive transcriptome data (Birzele et al. 2010; Rupp et al. 2014). In 2011, Becker et al. (2011) used NGS to unravel the transcriptome of CHO cells. They identified 29,184 isotigs (possible transcripts) and 24,576 isogroups (possible genes) from several CHO cell lines cultured under different conditions, and could assign 13,187 transcripts with functional annotation. From the transcriptome data, they reconstructed the central carbohydrate metabolism and the biosynthesis routes of sugars used for protein N-glycosylation, showing that the major steps of the N-glycosylation pathway are present in CHO cells.

In addition to mRNA, there is an increasing interest in non-coding RNAs especially microRNAs (miRNAs). Mature miRNAs are short non-coding RNAs (~22 nt) that regulate gene expression at a post-transcriptional level mainly by mRNA degradation or translational repression. The mature miRNA sequences are generally well conserved between animals, which allows the use of cross-species microarray platforms (human, mouse or rat). Hackl et al. (2012a) showed that >70% of the annotated CHO miRNAs can be reliable detected using microarrays comprising combined probe-sets against human, mouse and rat miRNAs. Hence, most of the published CHO miRNA profiling data were acquired using either human, mouse, rat or combining microarray platforms. But also NGS technologies (RNA-Seq) were applied (Hackl et al. 2011; Hammond et al. 2012b; Johnson et al. 2011) and could confirmed the high degree of miRNA sequence conservation between Chinese hamster, human, mouse and rat. In these three NGS studies, between 190 and 365 conserved mature miRNAs were identified in CHO cells.

To date, few miRNA profiling studies have been conducted in order to understand their function and relevance for specific phenotypes in CHO cells. Mainly microarrays and confirmation by qRT-PCR was applied to measure the mature miRNA transcript levels under certain conditions in order to relate them to cell growth/proliferation, apoptosis and q_P. In the first miRNA profiling study, Gammell et al. (2007) analyzed changes in miRNA expression during batch cultivation as well as upon low temperature shift (37°C to 31°C), which is a common cultivation strategy to achieve growth arrest that typically results in increased productivity and prolonged longevity. They found miR-21 and miR-24, two known growth inhibitory miRNAs, as being up-regulated during stationary growth phase induced either by the temperature shift or during normal batch culture. In another low temperature shift study, Barron et al. (2011) applied qRT-PCR (human TagMan microRNA array) to profile miRNA expression. In addition to several differentially expressed miRNAs, they found miR-7 more than 8-fold down-regulated upon low temperature shift. Overexpression of miR-7 confirmed its role as inhibitor of cell growth accompanied by increased q_P, however, down-regulation of miR-7 using antagonists did not increase growth of recombinant CHO cells. In another study, Hernandez Bort et al. (2012) analyzed the dynamics of miRNA expression during batch cultivation of CHO-K1 cells. In total, they identified 118 miRNAs that were significantly differentially expressed during batch cultivation relative to the lag phase. Special attention was given to the miR-17-92 cluster which was up-regulated during the exponential growth phase. The transient overexpression of miR-17 in an Epo-Fc secreting CHO cell line increased the growth rate significantly (Jadhav et al. 2012), and the stable overexpression of

miR-17 resulted in enhanced growth and a two-fold increased q_P (Jadhav et al. 2014). In an extensive study, Clarke et al. (2012) investigated the role of miRNAs in the regulation of CHO cell growth by profiling the miRNA expression of 30 clones with different specific growth rates ranging from 0.011 to 0.044 h^{-1}. They identified 51 miRNAs that were associated with increased growth rate. In a more recent study, Hackl et al. (2014) analyzed miRNA transcription of five CHO cell lines with low to high specific growth rate using microarrays. They identified 63 growth-correlated miRNAs, of which 43 were positively correlated. In addition, they found that the expression of Dicer, a key enzyme of the miRNA biogenesis which mediates the final step of miRNA maturation, was strongly correlated with cell growth.

To address the relevance of miRNAs for apoptosis in CHO cells, Druz et al. (2011) analyzed changes in miRNA transcription upon nutrient depletion using microarray technology. They identified 70 miRNAs that were significantly differentially expressed in the cells exposed to depleted media compared to cells exposed to fresh media. Among these miRNAs, the miR-297-669 cluster was strongly up-regulated in CHO cells subjected to depleted media and special attention was given to miR-466h-5p which was 452-fold induced. In a follow-up study, Druz et al. (2013) showed that the stable inhibition of mmu-miR-466h-5p improves apoptosis resistance and protein production in recombinant CHO cells.

To date, only two miRNA profiling studies have addressed the question whether miRNAs are directly involved in driving or maintaining a high q_P phenotype in recombinant CHO cells. Lin et al. (2011) compared the miRNA expression profiles of four IgG-producing cell lines to the parental DG44 host cell line using a cross-species microarray platform for screening and qRT-PCR for confirmation. They observed a significant down-regulation of miR-221 and miR-222 in all IgG-producing cell lines compared to the host, however, there was no significant correlation between q_P and the relative expression of these two miRNAs. Hackl et al. (2011) compared the miRNA expression of an Epo-Fc and an antibody producing cell line to the respective parental CHO cell line using RNA-Seq. They identified one miRNA (miR-21-5p) which was significantly down-regulated in both recombinant cell lines and seven miRNAs (miR-10b-5p, miR-16b-5p, miR-28-5p, miR-1959-5p, miR-125a-5p, miR-1937-5p and miR-1839-5p) which were significantly up-regulated. According to miRBase 20 (Kozomara and Griffiths-Jones 2014), two of them (miR-1959-5p and miR-1937-5p) are, however, fragments of tRNAs and therefore most likely no miRNAs.

In addition to miRNA, other non-coding RNAs such as long non-coding RNAs and PIWI-interacting RNAs (piRNAs) might be interesting for cell engineering of CHO cells (Jadhav et

al. 2013). In a first attempt to identify transcribed piRNAs in CHO cells, Gerstl et al. (2013) predicted 25,626 piRNAs and 540 piRNA clusters from RNA-Seq data.

1.6.2 Proteomics

Transcriptome analysis can provide first insights into the regulation of cellular processes, however, the expression levels of the gene products need not directly correlate with the mRNA levels. Proteomics is the detection and quantification of the complete set of proteins within a cell (proteome) under specific physiological conditions. Hence, proteomics provides more detailed information about the behavior of cells compared to mRNA expression profiling. However, the analysis of the proteome is analytically more challenging due to the wide dynamic range of protein expression levels (7–8 orders of magnitude) in mammalian cells (Anderson and Anderson 1998). The present methods are either gel or mass spectrometry (MS)-based. Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) allows the analysis of a large number of proteins separated by charge and mass. However, it is limited in its ability to detect proteins of low abundance, high hydrophobicity, low or high molecular weight, extreme isoelectric points or proteins that co-migrate (Rabilloud et al. 2010). Two-dimensional difference gel electrophoresis (2D-DIGE) enables the analysis and comparison of up to three samples (most commonly two samples and an internal standard) on one gel by labeling the samples with different fluorescent dyes, which improves the reproducibility and reliability of differential protein expression analysis (Alban et al. 2003). To overcome the limitations of gel-based techniques, shotgun proteomics approaches have been developed that couple high-performance liquid chromatography (HPLC) and tandem mass spectrometry (MS/MS) (Fournier et al. 2007). For relative quantification of protein levels between different cell samples, peptides can be labeled using techniques such as stable isotope labeling (SILAC), isobaric tags for relative and absolute quantification (iTRAQ) and tandem mass tag (TMT).

The recently available CHO genomic sequence data have greatly improved the capacity of MS-based proteomics studies by increasing the number of identifiable proteins as well as the confidence in identification (Meleady et al. 2012b). In a recent comprehensive study, the proteome of the CHO-K1 cell line was elucidated using different fractionation strategies and enrichment methods combined with MS analysis, leading to the identification of 6164 proteins (Baycin-Hizal et al. 2012). In addition, the codon frequency of CHO cells was determined and found to be distinct from humans.

Comparative proteomics in CHO cells has mainly been done to study the intracellular and physiological changes upon q_P enhancing treatments such as low culture temperature shift (Baik et al. 2006; Kantardjieff et al. 2010; Kaufmann et al. 1999; Kumar et al. 2008), hyperosmotic pressure (Lee et al. 2003), sodium butyrate (Baik et al. 2008; Kantardjieff et al. 2010; Yee et al. 2008a), cytochalasin D (Hayduk and Lee 2005) and dimethyl sulfoxide (DMSO) treatment (Li et al. 2006). Similarly, the proteomes of CHO cells with high μ and low μ were compared in order to identify proteins related to enhanced growth rate (Doolan et al. 2010). Comparative proteomics was also applied to study the intracellular effects of cell engineering including the overexpression of the apoptosis inhibitor Bcl-x_L (Baik and Lee 2010; Carlage et al. 2009; Carlage et al. 2012), the overexpression of PACEsol which effects q_P (Meleady et al. 2008) and the overexpression of miR-7 which significantly inhibits growth but enhances q_P (Meleady et al. 2012a). Only few studies have addressed the identification of cellular proteins that are differentially expressed between CHO clones exhibiting low and high q_P (Kang et al. 2014; Nissom et al. 2006).

Proteomics of CHO cells is frequently combined with transcriptome analysis in order to monitor transcriptional and post-transcriptional events (Baik et al. 2006; Doolan et al. 2010; Kang et al. 2014; Kantardjieff et al. 2010; Nissom et al. 2006; Yee et al. 2008a). In a recent study, Clarke et al. (2012) combined miRNA, mRNA and protein expression analysis in order to reveal the role of post-transcriptional regulation by miRNAs controlling cell growth of recombinant CHO cells.

1.6.3 Metabolomics

Transcriptomics and proteomics provide information about the presence and expression levels of cellular proteins, however, not about the biochemical reactions within a cell. The concentrations of metabolites are determined by the activity of many enzymes. Hence, metabolites are the direct signature of biochemical activity and therefore easier to correlate with the cellular phenotype (Patti et al. 2012). The metabolome comprises all small molecules produced by a cell including carbohydrates, amino acids, nucleosides, amines and fatty acids (Datta et al. 2013). Metabolomics aims to detect and quantify all of these molecules and is either based on nuclear magnetic resonance (NMR) or high performance mass spectrometry such as liquid chromatography mass spectrometry (LC-MS), gas chromatography mass spectrometry (GC-MS) and capillary electrophoresis mass spectrometry (CE-MS). However metabolomics faces significant technical challenges caused by rapid turnover,

physicochemical diversity and varying abundance of metabolites (Dietmair et al. 2010). To address the rapid turnover of certain metabolites, various quenching and extraction methods have been established and optimized in order to improve the preservation of the *in vivo* metabolite levels of mammalian cells (Dietmair et al. 2010; Hernandez Bort et al. 2014; Sellick et al. 2011b; Volmer et al. 2011).

Metabolomics profiling of extracellular metabolites has been used to design improved media and feeding strategies that enhance recombinant protein production in CHO cells (Bradley et al. 2010; Chong et al. 2009; Sellick et al. 2011a). Based on the identified nutrient limitations, Sellick et al. (2011a) developed a feeding strategy that significantly increased cell biomass and doubled the antibody titer. Chong et al. (2009) observed a significant extracellular accumulation of malate, which was caused by an enzymatic bottleneck at the malate dehydrogenase II (MDH II) in the tricarboxylic acid (TCA) cycle. Overexpression of MDH II increased the intracellular levels of ATP and NADH as well as the maximal viable cell concentration (Chong et al. 2010). Metabolomics was also used to identify apoptosis-inducing metabolites in the media of antibody-producing CHO fed-batch cultures (Chong et al. 2011). They found several metabolites, mainly nucleosides/nucleotides and amino acid derivatives, that correlated with caspase activity, where oxidized glutathione, AMP and GMP were shown to induce apoptosis when introduced to fresh CHO cultures.

The more challenging profiling of intracellular metabolites has already been done to identify metabolites associated with q_P (Chong et al. 2012) and growth rate (Dietmair et al. 2012). Chong et al. (2012) identified several metabolites (NADH, FAD, reduced and oxidized glutathione, and three activated sugar precursors) that are potentially related to the high q_P phenotype. These metabolites are involved in key cellular pathways of the TCA cycle, oxidative phosphorylation, glutathione metabolism and glycosylation.

2. Objectives

Recombinant Chinese hamster ovary (CHO) cells have now been used for more than 25 years as a host for the production of therapeutic proteins. Although the recent advances in cell culture technology have led to product titers up to 10 g L^{-1} in industry, the cellular processes that control and limit recombinant protein synthesis and secretion as well as cell growth are poorly understood. Systems biology approaches such as transcriptomics, proteomics and metabolomics can provide an important contribution to a comprehensive characterization of mammalian cell factories. However, the study of CHO cells has been hindered due to the lack of genomic sequence information. Hence, the publication of the first genomic sequence of a CHO cell line in 2011 was a major breakthrough which heralded a new era of CHO systems biology.

The present thesis is part of an extensive cross-species comparison between the two prominent expression systems CHO cells and *Pichia pastoris*. A prerequisite for a thorough and meaningful system comparison is the establishment of comparable recombinant CHO cell lines and *P. pastoris* strains as well as the cultivation in comparable bioreactor systems under comparable conditions. The present work comprises the development, cultivation and analysis of the recombinant CHO cell lines, whereas the study of *P. pastoris* is part of a separate PhD project.

The aims of this thesis were as follows:

- Development of stable low and high producing CHO cell lines expressing either human serum albumin (HSA) or the 3D6 anti-HIV-1 single-chain Fv-Fc fusion antibody (3D6scFv-Fc) by selecting clones at different stages of transgene amplification.
- Conducting a system comparison between the established high producing CHO cell lines and *P. pastoris* stains cultivated in comparable fed batch bioreactor processes, in order to highlight the strengths and limitations of each expression system.
- Cultivation of the established recombinant CHO cell lines as well as a mock transfected non-producer cell line in a bioreactor under steady-state conditions using a chemostat process, in order to generate comparable samples for transcriptomics (mRNA and microRNA) and proteomics.

- Comparative transcriptome (mRNAs) analysis of the established high, low and non-producing CHO cell lines using a novel CHO-specific microarray and qRT-PCR, in order to identify genes that are involved and potentially constitute bottlenecks in pathways and processes associated with high recombinant protein synthesis and secretion.
- Comparative microRNA expression profiling of the established high, low and non-producing CHO cell lines using a cross-species microarray and qRT-PCR, in order to identify microRNAs associated with a high productivity phenotype.

The outcome of this study enables the identification of genes and microRNAs that are potential targets for cell engineering in order to improve process-relevant phenotypes such as high specific secretion rate (q_P) or longevity.

In addition, this work represents the foundation for a genome-wide comparison of recombinant protein expression between CHO cells and *P. pastoris*, with the aim to identify organism-wide conserved and species-unique mechanisms and pathways that influence recombinant protein production.

3. Relevant publications

3.1 Publication A: Nuclear DNA content of CHO cells

Whole genome sequencing improves estimation of nuclear DNA content of Chinese hamster ovary cells

Andreas Maccani, Wolfgang Ernst and Reingard Grabherr

Published in Cytometry Part A (2013)

Cytometry A 83: 893-895

DOI: 10.1002/Cyto.A.22331

Cytometry
PART A
Journal of the International Society for Advancement of Cytometry

Whole Genome Sequencing Improves Estimation of Nuclear DNA Content of Chinese Hamster Ovary Cells

CHINESE hamster ovary (CHO) cells are the most widely used mammalian host for the expression of recombinant proteins for therapeutic application. The generation of stable CHO production cell lines as well as the methods for characterization and validation are well established. High-expressing cell lines are generally produced by exploiting the phenomenon of gene amplification which is widespread in eukaryotic cells (1). The host cells (e.g., dihydrofolate reductase-deficient CHO cells) are transfected with an expression vector containing an adequate amplification marker (e.g., dihydrofolate reductase) and the gene of interest. Applying increasing amounts of the appropriate inhibitor (e.g., methotrexate) leads to the amplification of the amplification marker as well as genetically linked sequences including the gene of interest (Fig. 1). In order to study the efficiency of the applied amplification strategy, the correlation between transgene copy number and target transcript or target protein levels, or the stability of the copy number over cultivation time, it is essential to know the number of target genes that are integrated in the genome. The method most frequently used to determine gene copy numbers is quantitative real-time PCR (qPCR). The calculation of the absolute gene copy number per cell requires the knowledge of the number of genome copies in the template used for amplification. However, the assumed nuclear DNA content of CHO cells varies considerably between different studies ranging from 3.3 pg (2) to 6.0 pg per cell (3) or is not mentioned at all in many other publications. Although the sequence of the CHO-K1 cell line became available recently (4), there is still some uncertainty among scientists about the nuclear DNA amount in unreplicated nuclei of CHO cells. The diploid genome of a normal Chinese hamster (*Cricetulus griseus*) cell has 22 chromosomes (10 pairs of autosomes and 2 sex chromosomes). However, CHO cells have an altered karyotype. In the parental cell line of CHO-K1, only 8 of the determined 21 chromosomes appear to be normal as compared to the euploid Chinese hamster chromosomes and the remaining 13 chromosomes (Z group) contain translocations, deletions, and pericentric inversions (5). Nevertheless, apart from the X_2 chromosome all of the euploid chromatin is present in CHO cells (5). Consequently, CHO cells have a nuclear DNA content close to that of normal diploid Chinese hamster cells.

Furthermore, the definition of genome size often leads to confusion. The correct usage of the term "genome size" and also the commonly used term "C-value" has been discussed extensively in literature (6). For diploid cells, genome size and C-value are interchangeable and are generally defined as the amount of nuclear DNA within a haploid ($1n$) set of chromosomes (7,8). Consequently, the nuclear DNA content of a diploid cell is twice the genome size or the 2C-value. However, what does this mean for CHO cells? Here, the unreplicated nuclear DNA content of CHO cells (2C DNA amount) was estimated in three different ways based on the currently available data from literature providing very similar results.

THE GENOME SEQUENCE OF THE CHO-K1 CELL LINE

In 2011, the genome sequence of the CHO-K1 cell line was published (4). The assembly contained 2.45 Gb of genomic sequence and the authors of this study estimated a genome size of 2.6 Gb (4). As the total amount of chromatin in CHO cells is close to that in normal Chinese hamster cells (5), it can be assumed that most of the genomic sequence is also present twice in CHO cells like it would be in normal

Received 5 April 2013; Revised 14 May 2013; Accepted 7 June 2013

Grant sponsors: Federal Ministry of Economy, Family and Youth (BMWFJ), Federal Ministry of Traffic, Innovation and Technology (bmvit), Styrian Business Promotion Agency SFG, Standortagentur Tirol and ZIT—Technology Agency of the City of Vienna through the COMET-Funding Program managed by the Austrian Research Promotion Agency FFG.

*Correspondence to: Andreas Maccani, Austrian Centre of Industrial Biotechnology (ACIB GmbH), Muthgasse 11, 1190 Vienna, Austria.
E-mail: andreas.maccani@acib.at

Published online 10 July 2013 in Wiley Online Library (wileyonlinelibrary.com)

DOI: 10.1002/cyto.a.22331

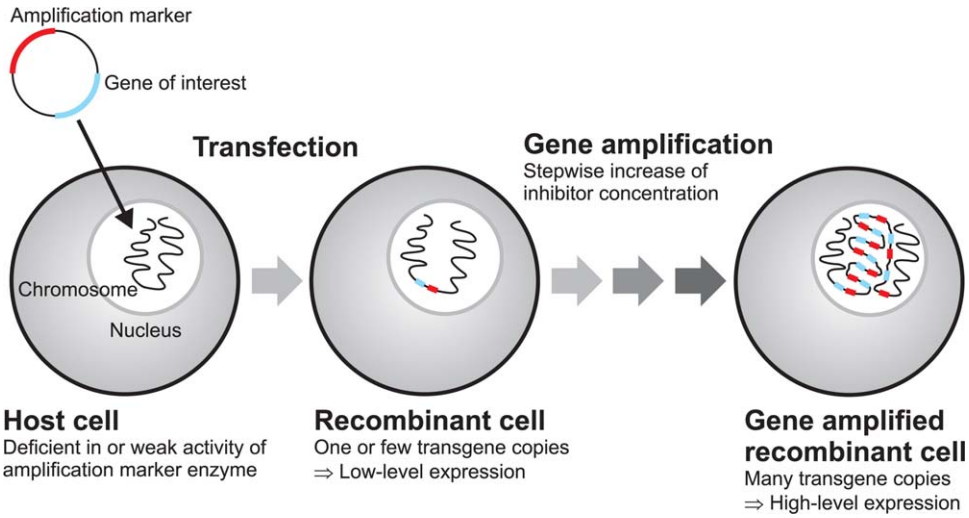


Figure 1. General procedure of gene amplification to generate stable, high-expressing recombinant mammalian cell lines. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

diploid cells. Consequently, the 2C genome size is 5.2 Gb. Applying Eq. (1) which was determined by Doležel et al. (9) leads to a 2C DNA amount of 5.3 pg.

$$\text{DNA content}[\text{pg}] = \text{genome size}[\text{bp}] / (0.978 \times 10^9) \quad (1)$$

Feulgen Cytophotometry

In a study from 1983, Feulgen cytophotometry was used to determine the nuclear DNA content of various animal cells relative to the plant *Allium cepa* (10). *Allium cepa* is frequently used as reference for DNA measurements having a 2C-value of 33.5 pg (11). Although this value originates from a study published in 1965, it is still the generally accepted estimate today (7). Greilhuber et al. measured 134 nuclei of mitotic Chinese hamster lung fibroblasts and 50 nuclei of mitotic CHO cells which were considered as diploid ($2n$). The determined DNA amount relative to *Allium cepa* was 0.1590 ± 0.0088 for the Chinese hamster cells and 0.1630 ± 0.0058 for CHO cells (10). Consequently, the 2C DNA amounts are 5.33 pg and 5.46 pg, respectively.

Comparison to the Human Genome Size

The overall human genome (22 autosomes, X and Y chromosome) was estimated to be 3.08 Gb (12). Subtracting the size of the Y chromosome leads to a genome size of 3.05 Gb for a female human cell. Kato et al. measured the genome size of 34 species relative to diploid human fibroblasts using Feulgen cytophotometry (13). They determined the genome size of the Chinese hamster to be 91.5% of the female human genome. Hence, the genome size of the Chinese hamster is 2.79 Gb. As CHO cells contain 3% less total chromatin than diploid Chinese hamster cells (5), the 2C genome size of CHO cells is 5.41 Gb. Applying Eq. (1) leads to a 2C DNA amount of 5.53 pg.

The human genome size of 3.05 Gb corresponds to a 1C-value of 3.12 pg which perfectly fits to the 3.11 ± 0.16 pg measured previously by Greilhuber et al. using Feulgen cytophotometry (10). They determined the nuclear DNA content of CHO cells to be 88% of human, which would lead to a 2C DNA amount of 5.49 pg.

Average Nuclear DNA Content of CHO Cells

All considerations were made based on data from the CHO-K1 cell line or its parental cell line. CHO cells are genetically very unstable leading to considerable genomic heterogeneity between different cells lines (14). For example, the cell line DG44 has only 20 chromosomes of which only seven are normal, four can be assigned to the Z group, and the remaining nine are altered chromosomes (15). Consequently, it is very likely that there are also slight differences in genome size between cell lines. However, it can be assumed that the variations in the total amount of nuclear DNA are marginal. The determined 2C DNA amounts of CHO cells were within a narrow range of 5.3–5.53 pg. For this reason 5.4 pg seems to be a good estimation of the unreplicated nuclear DNA content of CHO cells.

Andreas Maccani*
Austrian Centre of Industrial Biotechnology (ACIB GmbH)
Muthgasse 11, 1190 Vienna, Austria

Wolfgang Ernst[1,2]
[1]Austrian Centre of Industrial Biotechnology (ACIB GmbH)
Muthgasse 11, 1190 Vienna, Austria

[2]Department of Biotechnology
University of Natural Resources and Life Sciences
Muthgasse 18, 1190 Vienna, Austria

Reingard Grabherr
Department of Biotechnology
University of Natural Resources and Life Sciences
Muthgasse 18, 1190 Vienna, Austria

Literature Cited

1. Omasa T. Gene amplification and its application in cell and tissue engineering. J Biosci Bioeng 2002;94:600–605.
2. Reisinger H, Steinfellner W, Stern B, Katinger H, Kunert R. The absence of effect of gene copy number and mRNA level on the amount of mAb secretion from mammalian cells. Appl Microbiol Biotechnol 2008;81:701–710.
3. Osterlehner A, Simmeth S, Gopfert U. Promoter methylation and transgene copy numbers predict unstable protein production in recombinant Chinese hamster ovary cell lines. Biotechnol Bioeng 2011;108:2670–2681.
4. Xu X, Nagarajan H, Lewis NE, Pan S, Cai Z, Liu X, Chen W, Xie M, Wang W, Hammond S, et al. The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. Nat Biotechnol 2011;29:735–741.
5. Deaven LL, Petersen DF. The chromosomes of CHO, an aneuploid Chinese hamster cell line: G-band, C-band, and autoradiographic analyses. Chromosoma 1973;41: 129–144.
6. Greilhuber J, Doležel J, Lysák MA, Bennett MD. The origin, evolution and proposed stabilization of the terms 'genome size' and 'C-value' to describe nuclear DNA contents. Ann Bot 2005;95:255–260.
7. Doležel J, Greilhuber J. Nuclear genome size: Are we getting closer? Cytometry A 2010;77A:635–642.
8. Gregory TR. Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. Biol Rev Camb Philos Soc 2001;76:65–101.
9. Doležel J, Bartoš J, Voglmayr H, Greilhuber J. Nuclear DNA content and genome size of trout and human. Cytometry A 2003;51A:127–128.
10. Greilhuber J, Volleth M, Loidl J. Genome size of man and animals relative to the plant Allium cepa. Can J Genet Cytol 1983;25:554–560.
11. Van't Hof J. Relationships between mitotic cycle duration, S period duration and the average rate of DNA synthesis in the root meristem cells of several plants. Exp Cell Res 1965;39:48–58.
12. International Human Genome Sequencing C. Finishing the euchromatic sequence of the human genome. Nature 2004;431:931–945.
13. Kato H, Harada M, Tsuchiya K, Moriwaki K. Absence of correlation between DNA repair in ultraviolet irradiated mammalian cells and life span of the donor species. Jpn J Genetics 1980;55:99–108.
14. Cao Y, Kimura S, Itoi T, Honda K, Ohtake H, Omasa T. Construction of BAC-based physical map and analysis of chromosome rearrangement in Chinese hamster ovary cell lines. Biotechnol Bioeng 2012;109:1357–1367.
15. Derouazi M, Martinet D, Besuchet Schmutz N, Flaction R, Wicht M, Bertschinger M, Hacker DL, Beckmann JS, Wurm FM. Genetic characterization of CHO production host DG44 and derivative recombinant cell lines. Biochem Biophys Res Commun 2006;340:1069–1077.

3.2 Publication B: Bioprocess comparison – CHO cells vs. *Pichia pastoris*

Pichia pastoris secretes recombinant proteins less efficiently than Chinese hamster ovary cells but allows higher space-time yields for less complex proteins

Andreas Maccani, Nils Landes, Gerhard Stadlmayr, Daniel Maresch, Christian Leitner, Michael Maurer, Brigitte Gasser, Wolfgang Ernst, Renate Kunert and Diethard Mattanovich

Published in Biotechnology Journal (2014)

Biotechnology Journal 9: 526-537

DOI:10.1002/biot.201300305

Biotechnology Journal
www.biotechnology-journal.com

Biotechnol. J. 2014, 9, 526–537

DOI 10.1002/biot.201300305

Research Article

Pichia pastoris secretes recombinant proteins less efficiently than Chinese hamster ovary cells but allows higher space-time yields for less complex proteins

Andreas Maccani[1,2,], Nils Landes[1,2,*], Gerhard Stadlmayr[3,4], Daniel Maresch[4], Christian Leitner[2], Michael Maurer[1,5], Brigitte Gasser[1,2], Wolfgang Ernst[1,2], Renate Kunert[1,2] and Diethard Mattanovich[1,2]*

[1] Austrian Centre of Industrial Biotechnology (ACIB GmbH), Vienna, Austria
[2] Department of Biotechnology, University of Natural Resources and Life Sciences, Vienna, Austria
[3] Christian Doppler Laboratory for Antibody Engineering, University of Natural Resources and Life Sciences, Vienna, Austria
[4] Department of Chemistry, University of Natural Resources and Life Sciences, Vienna, Austria
[5] School of Bioengineering, University of Applied Sciences FH-Campus Vienna, Vienna, Austria

Chinese hamster ovary (CHO) cells are currently the workhorse of the biopharmaceutical industry. However, yeasts such as *Pichia pastoris* are about to enter this field. To compare their capability for recombinant protein secretion, *P. pastoris* strains and CHO cell lines producing human serum albumin (HSA) and the 3D6 single chain Fv-Fc anti-HIV-1 antibody (3D6scFv-Fc) were cultivated in comparable fed batch processes. In *P. pastoris*, the mean biomass-specific secretion rate (q_P) was 40-fold lower for 3D6scFv-Fc compared to HSA. On the contrary, q_P was similar for both proteins in CHO cells. When comparing both organisms, the mean q_P of the CHO cell lines was 1011-fold higher for 3D6scFv-Fc and 26-fold higher for HSA. Due to the low q_P of the 3D6scFv-Fc producing strain, the space-time yield (STY) was 9.6-fold lower for *P. pastoris*. In contrast, the STY of the HSA producer was 9.2-fold higher compared to CHO cells because of the shorter process time and higher biomass density. The results indicate that the protein secretion machinery of *P. pastoris* is much less efficient and the secretion rate strongly depends on the complexity of the recombinant protein. However, process efficiency of the yeast system allows higher STYs for less complex proteins.

Received	17 JUL 2013
Revised	22 NOV 2013
Accepted	27 DEC 2013
Accepted article online	03 JAN 2014

Supporting information available online

Keywords: CHO cells · Fed batch · Protein secretion · Recombinant protein production · Volumetric productivity

1 Introduction

Human protein therapeutics became more and more important for the treatment of various diseases over the last decades. Today, many different production systems are in use for the expression of heterologous proteins ranging from bacterial hosts to transgenic animals. To date, approved biopharmaceutical products are produced in a limited number of expression systems (in particular *Escherichia coli*, *Saccharomyces cerevisiae*, and Chinese hamster ovary (CHO) cells) [1], but non-conventional systems are catching up [2]. The choice of the most suitable expression system strongly depends on the complexity of the product as well as the need for correct post-translational modifications (PTMs) such as glycosylation, disulfide bond formation, phosphorylation, and proteolytic processing which might be required for biological efficacy.

Correspondence: Prof. Diethard Mattanovich, Department of Biotechnology, University of Natural Resources and Life Sciences, Muthgasse 18, 1190 Vienna, Austria
E-mail: diethard.mattanovich@boku.ac.at

Abbreviations: 3D6scFv-Fc, 3D6 single chain Fv-Fc anti-HIV-1 antibody; **CDM**, cell dry mass; **CHO**, Chinese hamster ovary; **HSA,** human serum albumin; **MTX,** methotrexate; **PTM,** post-translational modification; **STY**, space-time yield

* These authors contributed equally to this work.

Mammalian cells are currently the expression system of choice for the production of complex therapeutic proteins which require proper folding, human-like PTMs or multimeric assembly. Various mammalian cell lines such as CHO, mouse myeloma (NS0), baby hamster kidney (BHK), human embryonic kidney (HEK-293), and human embryonic retinoblast (PER.C6) cells have been established. However, almost all approved mammalian cell-derived biopharmaceutical proteins have been produced in CHO cells [1]. Mammalian cells have been used for more than three decades for the production of recombinant proteins and initially they were considered as the least effective production system reaching product titers of only 50 mg L^{-1} in the mid-1980s. But mainly due to media and bioprocess optimization, product titers ranging from 1 to 5 g L^{-1} are typically reached in industry today [3]. Compared to microbial expression systems, mammalian cell cultures grow very slowly and reach only a low biomass density. Moreover, media costs are generally higher, although chemically defined, serum- and protein-free media have been developed for CHO cells. Additionally, cell line development is a very time-consuming process which takes at least 6 months [4]. In spite of these drawbacks, mammalian cells and especially CHO cells are currently the workhorse in biopharmaceutical industry, mainly because they are the only established expression system which is able to produce complex recombinant proteins with human-like glycoforms which are bioactive in humans. But also yeast expression systems have been used for the production of therapeutic proteins since the early 1980s. They can grow on inexpensive, chemically defined media to very high cell densities. Until 2009, all approved yeast-derived biopharmaceutical proteins were expressed in *S. cerevisiae*. However, expression systems based on non-conventional yeasts have been developed during the last two decades. Especially the methylotrophic yeasts *Pichia pastoris* and *Hansenula polymorpha* are frequently applied for recombinant protein production. In 2009, Ecallantide (trade name Kalbitor) was the first therapeutic derived from *P. pastoris* that gained FDA approval [5]. The main advantages of yeasts over bacterial expression systems such as *E. coli* are their ability to secrete recombinant proteins into the culture broth as well as the absence of endotoxins. Moreover, yeasts do not contain oncogenic or viral DNA [6]. Furthermore, yeasts are eukaryotes and so they have the capability of protein processing such as protein folding and PTM-like glycosylation. *N*-linked glycosylation patterns however differ significantly from human and can impact the serum half-life and immunogenicity of therapeutic proteins [7, 8]. The high mannose content and the absence of sugar residues typical for human glycoproteins such as sialic acid are a major drawback of yeasts compared to mammalian expression systems. Because of this, much effort has been put into the engineering of strains that are able to form human-like glycans. The synthesis of fully humanized *N*-glycans in *P. pastoris* has been achieved [9], as well as the production of recombinant human proteins with humanized *N*-glycans [10].

Although a huge amount of information about the individual expression systems as well as comparative reviews can be found in literature, a quantitative comparison of heterologous protein production data is very difficult. The productivities of the individual systems strongly depend on the expressed recombinant protein. However, the proteins produced in microbial processes are generally different ones than those expressed in mammalian cells. Smaller proteins are supposed to be produced more economically in microbial cells, whereas mammalian expression systems are exclusively used for large glycosylated proteins. However, due to the recent advances in glycoengineering, yeasts become a more and more attractive alternative for the recombinant production of complex proteins.

In this study, process relevant parameters of high producing recombinant *P. pastoris* strains and CHO cell lines secreting the same model proteins were compared. For downstream processing, the product concentration as well as the relative purity of the culture supernatant is of high importance. Beside media costs, the achievable space-time yield (STY) is the crucial criterion to assess the economic efficiency of the fermentation process. The STY depends on the one hand on the specific growth rate (μ) and the achievable biomass density and on the other hand on the ability for product formation and secretion which is described by the specific product secretion rate (q_P). Two model proteins with different complexity were selected in order to challenge the expression systems in different ways. One of them, human serum albumin (HSA) is a monomeric and non-glycosylated protein that can be produced at very high levels in *P. pastoris* [11]. As a second more complex model protein, a single chain Fv-Fc fusion antibody (3D6scFv-Fc) derived from the monoclonal anti-HIV-1 antibody 3D6 [12] was designed. This protein is homodimeric and contains the Fc-specific glycosylation site. For both host systems, transgene copy number was increased by gene amplification in order to establish high producing strains and cell lines which then were cultivated in standard fed batch processes using the same bioreactor system. Comparing the process relevant parameters highlighted the strengths and limitations of *P. pastoris* and CHO cells for the production of recombinant proteins.

2 Materials and methods

2.1 Model protein construction

The 3D6scFv-Fc antibody was designed by combining the variable heavy chain (V_H) and the variable light chain (V_L) domain of the monoclonal antibody 3D6 [12] via a

$(GGGGS)_3$ linker and fusing this single-chain fragment variable (scFv) construct to the human IgG1 fragment crystallizable (Fc) region. The cDNAs of 3D6scFv-Fc and human serum albumin (HSA) were codon optimized for CHO cells and *P. pastoris* respectively and synthesized (Geneart, Germany).

2.1.1 *P. pastoris expression vector*

For both proteins, codon optimized genes were cloned into the multiple cloning site (SbfI, SfiI) of the in-house vector pPUZZLE containing the Zeocin resistance cassette for selection and the NTS region of the ribosomal DNA locus as genome integration sequence [13]. The expression of both model proteins was controlled by the constitutive glyceraldehyde-3-phosphate dehydrogenase (GAP) promoter of *P. pastoris*. HSA was secreted by means of its native secretion leader. For secretion of the 3D6scFv-Fc antibody the prepro leader sequence of the *S. cerevisiae* alpha mating factor was used.

2.1.2 *CHO cells expression vectors*

Both target genes were cloned into the pCI-neo mammalian expression vector (Promega, WI, USA) which carries the human cytomegalovirus (CMV) immediate-early enhancer/promoter to drive the constitutive expression of the inserted gene as well as the neomycin phosphotransferase gene for selection (pCI-neo_HSA_CHO, pCI-neo_3D6scFc_CHO). For secretion of HSA the native leader was used. The 3D6scFv-Fc antibody was secreted using the human Ig heavy chain leader. Additionally, a second plasmid (p2-dhfr) which contains the dihydrofolate reductase gene under the control of the SV40 early promoter was used for gene amplification.

2.2 *P. pastoris* strains and CHO cell lines

2.2.1 *P. pastoris strain development*

The establishment of a high producing *P. pastoris* strain for each model protein was based on the procedure of post-transformational vector amplification via repeated selection on stepwise increased antibiotic concentrations as described previously [14].

Plasmids linearized with *Spe*I were transformed into *P. pastoris* SMD1168H (Life Technologies, CA, USA) using electroporation (2 kV, 4 ms, GenePulser, Bio-Rad, CA, USA). After regeneration, the cell suspension was plated on YPD agar plates (10 g L^{-1} yeast extract, 10 g L^{-1} peptone, 20 g L^{-1} glucose, and 20 g L^{-1} agar) containing 25 μg mL^{-1} Zeocin. Initially, 24 clones for each model protein were picked from the 25 μg mL^{-1} Zeocin containing YPD agar plates, screened in shake flasks and analyzed by SDS–PAGE, western blot, and ELISA. Out of those, the best 12 clones were stepwise transferred to YPD agar plates with increasing Zeocin concentrations (100, 500, 1000, 2500, and 5000 μg mL^{-1}). Thus, 12 clone families were generated, each one containing six clones which were descended from different Zeocin levels. Thereby, the clone selected on the lower Zeocin level represents the parental strain of the clone selected on the next higher level. Screening of the corresponding clones was carried out in shake flask cultures on a Multitron II shaker (Infors, Switzerland). Therefore, a single colony of the desired clones was cultivated in 5 mL of YPD (10 g L^{-1} yeast extract, 10 g L^{-1} peptone, and 20 g L^{-1} glucose) medium supplemented with the respective amount of Zeocin. Such pre-cultures were shaken at 180 rpm for 24–48 h at 28°C. After measuring the optical density (OD_{600}) of the pre-cultures, main cultures (10 mL YPD medium in a 100 mL shake flask) were inoculated to an OD_{600} of 0.1 and grown for 48 h at 28°C and 180 rpm. Additional glucose (100 μL of 50% w/v glucose) was added to the cultures after 12, 24, and 36 h. The cultures were harvested after 48 h of cultivation. Wet cell mass concentrations were determined by centrifugation of 1 mL culture broth for 1 min at 17 000*g* and 4°C. Aliquots of the supernatant as well as cell pellets were stored at –20°C until further analysis.

2.2.2 *CHO cell line development*

Two stable recombinant CHO cell lines producing 3D6scFv-Fc and HSA, respectively were established. Protein-free cultivated dihydrofolate reductase deficient ($dhfr^-$) CHO cells DUKX-B11, ATCC CRL-9096 [15] were used as host cell line. These cells were co-transfected with the plasmids pCI-neo_HSA_CHO or pCI-neo_3D6scFc_CHO and p2-dhfr using polyethyleneimine (PEI) as previously described [16]. Selection of recombinant cell lines was performed in the presence of G418 and the absence of hypoxanthine and thymidine (HT). For this purpose the protein-free CHO medium ProCHO5 (Lonza, Switzerland) supplemented with 4 mM L-glutamine and 0.5 mg mL^{-1} G418 was used. To select single clones, limiting dilution was conducted by seeding the cells into 96-well plates for 24 h after transfection. For each cell line, transfections were done in four independent experiments using 5×10^6 cells respectively. Cells were cultivated in a 37°C, 5% carbon dioxide environment. Additionally, 0.05 μM methotrexate (MTX) was added to the medium for gene amplification and stepwise increased to 0.1 μM in the subsequent passages. Best producing clones were identified by screening the supernatants using product specific ELISA assays. The number of clones was stepwise reduced and the culture volume increased to 10 mL using 48-well plates and T25 cell culture flasks. Finally, the four best producing clones were transferred into 125 mL spinner flasks and propagated in 50 mL suspension cultures at 50 rpm and 37°C. These clones were evaluated regarding specific growth rate and specific productivity for at least ten passages. The best performing clone was then subcloned by limiting dilution in 96-well plates. The cultures were treated as before and MTX concentration was stepwise increased to 0.4 μM in the subsequent pas-

sages. The four best producing clones were again transferred into 125 mL spinner flasks and evaluated regarding specific growth rate and specific productivity for at least ten passages. For the two best performing clones, the MTX-pressure was stepwise increased to 0.8 μM and finally to 1.6 μM. The best performing clone at 0.4 μM MTX was subcloned by limiting dilution in 96-well plates for a third time. During culture volume expansion, the clones were adapted to DMEM/Ham's F12 (1:1) supplemented with 4 mM L-glutamine, 0.25% soy peptone (Quest International, The Netherlands), 0.1% Pluronic F68, protein-free supplement (Polymun Scientific, Austria), and 0.4 μM MTX. The four best producing clones were again transferred into 125 mL spinner flasks and evaluated regarding specific growth rate and specific productivity for at least ten passages. Working cell banks were generated at the different stages of gene amplification and subcloning.

2.3 Fed batch cultivation

2.3.1 *P. pastoris* cultivation

Glucose limited fed batch cultivations of the selected *P. pastoris* high producing strains were carried out in duplicate in 1.0 L bioreactors (SR0700ODLS, DASGIP, Germany) with a fed batch starting volume of 350 mL as described previously [17]. Pre-cultures for fed batch experiments were inoculated from a cryo-stock and grown at 25°C in 1000 mL shake flasks containing 100 mL of YPD medium with corresponding amounts of Zeocin. Cultures were shaken at 180 rpm for 24–48 h. Pre-culture cells were harvested and used to inoculate the bioreactor to the desired starting optical density (OD_{600}) of 3.0. Therefore, a defined culture broth volume was withdrawn, transferred into sterile 50 mL centrifuge tubes and centrifuged (1504*g*, 4 min, 25°C). After washing, the cells were re-suspended in 40 mL of sterile batch medium and used to inoculate the bioreactor, which was prefilled with 410 mL of sterile batch medium. After complete consumption of glycerol in the batch phase, fed batch cultivation was initiated by starting the balance controlled feed pumps. A constant feed of 2.38 g h^{-1} fed batch medium was applied for 113 h. The air flow rate was 13.5 L h^{-1}. Foam formation was antagonized by controlled addition of 5% w/w antifoam solution (Glanapon 2000, Bussetti, Austria). Samples for determination of cell dry mass (CDM), product concentrations and cell viability were taken every day. Viability of cells was analyzed with a cell viability kit (BD Biosciences, CA, USA) on a flow cytometer as described previously [18].

Batch medium contained per liter: 40 g glycerol, 12.6 g $(NH_4)_2HPO_4$, 0.5 g $MgSO_4 \cdot 7H_2O$, 0.9 g KCl, 0.022 g $CaCl_2 \cdot 2H_2O$, 2 g citric acid monohydrate, 4.6 mL PTM1 trace salts stock solution, and 2 mL of a 0.2 g L^{-1} biotin (Sigma–Aldrich, MO, USA) solution. The pH was set to 5.85 with 25% w/w HCl. Fed batch medium contained per liter: 550 g glucose·H_2O, 10 g KCl, 6.45 g $MgSO_4 \cdot 7H_2O$, 0.35 g $CaCl_2 \cdot 2H_2O$, 12 mL PTM1 trace salts stock solution, and 2 mL of a 0.2 g L^{-1} biotin solution. PTM1 trace salts stock solution contained per liter: 6.0 g $CuSO_4 \cdot 5H_2O$, 0.08 g NaI, 3.36 g $MnSO_4 \cdot H_2O$, 0.2 g $Na_2MoO_4 \cdot 2H_2O$, 0.02 g H_3BO_3, 0.82 g $CoCl_2 \cdot 6H_2O$, 20.0 g $ZnCl_2$, 65.0 g $FeSO_4 \cdot 7H_2O$, and 5.0 mL H_2SO_4 (95–98% w/w).

2.3.2 CHO cell cultivation

Fed batch cultivation was conducted in 800 mL cell culture bioreactors (DS0700TPSS, DASGIP, Germany) with a starting working volume of 500 mL. The inocula were expanded in spinner flasks starting from the working cell bank. Exponentially growing cells from passage six were used for inoculation. The initial cell concentration was 2.5×10^5 cells mL^{-1}. The cultures were maintained at 37°C, pH 7.0, 30% dissolved oxygen, and an agitation speed of 80 rpm. The initial batch medium was DMEM/Ham's F12 (1:1) supplemented with 4 mM L-glutamine, 0.25% soy peptone (Quest International), 0.1% Pluronic F68, and protein-free supplement (Polymun Scientific). The feed medium for the 3D6scFv-Fc expressing cell line contained 20 g L^{-1} D-glucose, 28 mM L-glutamine, 5× MEM amino acids (PAA, Austria), and 10× MEM non essential amino acids (PAA). The feed medium for the HSA producing cell line was composed of 10 g L^{-1} D-glucose, 14 mM L-glutamine, 2.5× MEM amino acids, and 5× MEM nonessential amino acids. Both feed media were based on DMEM/Ham's F12 (1:1) supplemented with 0.25% soy peptone, 0.1% Pluronic F68, and protein-free supplement.

After 50 h batch cultivation, the cultures were continuously fed at a constant feed rate. The feed rate was adjusted daily to maintain the glucose concentration of the culture broth within the target range of 1.0–2.5 g L^{-1}. Samples for off-line analyses were taken once a day. D-Glucose, L-glutamine, L-glutamate, and ammonium concentrations were measured with a bioprofile analyzer (BioProfile 100 Plus, Nova Biomedical, MA, USA).

2.4 Biomass concentration determination

2.4.1 *P. pastoris*

Optical density (OD) of *P. pastoris* cultures was measured at a wavelength of 600 nm after dilution in ddH_2O. Yeast dry mass concentrations were determined in duplicate for each sample. Three milliliters of culture broth were centrifuged at 4307*g* for 5 min, pellets were washed with 5 mL ddH_2O, centrifuged and re-suspended in ddH_2O. This suspension was transferred to a pre-weighed beaker and dried at 105°C for 24 h.

2.4.2 CHO cells

Cell concentration was determined by counting the nuclei of lysed cells with a Z2 Coulter Counter (Beckman

Coulter, CA, USA). Cell viability was determined by trypan blue exclusion using a hemocytometer.

For CDM determination, more than 9×10^7 cells were collected by centrifugation, washed with Dulbecco's PBS and dried as described above. The CDM was determined once per fermentation. The time course of the CDM concentration was calculated based on the cell concentration and the determined average dry mass of one CHO cell.

2.5 Analytical methods

2.5.1 ELISA

The concentrations of the secreted products were determined from the culture supernatants using sandwich ELISA assays. For 3D6scFv-Fc, 96-well immunosorbent plates (Nunc MaxiSorp, Thermo Fisher Scientific, MA, USA) were coated with 0.33 μg mL^{-1} goat anti-human IgG (γ-chain specific) antibody (I3382, Sigma-Aldrich) diluted in coating buffer (0.1 M Na_2CO_3/$NaHCO_3$, pH 9.6) at 4°C overnight. After each incubation step, the plates were washed three times using washing buffer (phosphate-buffered saline (PBS)) containing 0.1% Tween 20, pH 7.4). Affinity purified 3D6scFv-Fc was used as a standard protein at a starting concentration of 100 ng mL^{-1}. Standard and samples were serially diluted in washing buffer containing 1% bovine serum albumin (BSA) and applied onto the pre-coated plates. After 1 h, captured 3D6scFv-Fc was incubated with 0.5 μg mL^{-1} horseradish peroxidase (HRP) conjugated goat anti-human IgG (γ-chain specific) antibody (62-8420, Life Technologies) for 1 h. Staining was conducted using 100 mg mL^{-1} o-phenylenediamine diluted in 0.15 M citric acid buffer, pH 5.0 containing 0.02% H_2O_2. After stopping the reaction with 25% H_2SO_4, absorption was measured at 492 nm (620 nm reference wavelength) using the infinite M1000 microplate reader (Tecan, Switzerland). HSA concentrations were determined using the Human Albumin ELISA Quantification Set (E80-129, Bethyl, TX, USA) according to the manufacturer's instructions.

2.5.2 SDS–PAGE and western blot

Sample supernatants were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) using NuPAGE Novex 4–12% Bis-Tris Gels (Life Technologies) and MOPS buffer according to the manufacturer's instructions. Proteins were visualized by silver staining [19] or transferred to a polyvinylidene difluoride (PVDF) membrane for western blot analysis. Semi wet blotting was applied using the XCell II Blot Module (Life Technologies) according to the supplier's manual. The membrane was blocked with washing buffer (PBS, 0.1% Tween 20) containing 3% skim milk powder for 1 h. To detect 3D6scFv-Fc, the membrane was incubated with alkaline phosphatase (AP) conjugated goat anti-human IgG (γ-chain specific) antibody (A3187, Sigma-Aldrich) 1:5000 diluted in washing buffer containing 3% skim milk powder. 5-Bromo-4-chloro-3-indolyl-phosphate combined with nitro blue tetrazolium (BCIP/NTB) staining was used for the colorimetric detection of AP activity. HSA was detected using a horseradish peroxidase (HRP) conjugated goat anti-human albumin antibody (A80-129P, Bethyl) in a 1:30 000 dilution. For protein visualization the Super-Signal West Pico Chemiluminescent Substrate (Thermo Fisher Scientific) was used.

2.6 Parameter calculations

Smoothing of experimental data for CDM and secreted product concentration was done using the smoothing spline algorithm of the Matlab Curve Fitting Toolbox (The MathWorks, MA, USA). A smoothing parameter of $p = 0.2$ was used for smoothing the CDM data, whereas a smoothing parameter of $p = 0.00001$ was applied for smoothing of the product concentration data. Specific secretion rates and STYs were calculated based on smoothed product concentrations as described in the following. The specific product secretion rate q_P (mg g^{-1} h^{-1}) between two consecutive sampling points was calculated according to Equation (1), where t_i (h) is the later point in time. P_i (mg) is the total amount of secreted product at the time t_i and CDM_i (g) represents the total CDM at the time t_i

$$q_P = \frac{1}{t_i - t_{i-1}} \frac{2(P_i - P_{i-1})}{CDM_{i-1} + CDM_i} \quad (1)$$

The mean specific product secretion rate $q_{P,mean}$ (mg g^{-1} h^{-1}) was calculated with Equation (2), where t_E (h) is the time at the end of cultivation. In this study t_0 (h) represents the time of the feed start because $q_{P,mean}$ was determined for the feed phase

$$q_{P,mean} = \frac{1}{t_E - t_0} \sum_{i-1}^{E} \frac{2(P_i - P_{i-1})}{CDM_{i-1} + CDM_i} \quad (2)$$

The STY (mg L^{-1} h^{-1}) at a given point t_i (h) was calculated according to Equation (3), where t_0 (h) represents the time of fermentation start. P_0 (mg) is the total amount of product at the time t_0. V_i is the total culture volume at the time t_i

$$STY = \frac{1}{V_i} \frac{P_i - P_0}{t_i - t_0} \quad (3)$$

The specific growth rate μ (h^{-1}) between two consecutive sampling points was calculated according to Equation (4)

$$\mu = \frac{1}{t_i - t_{i-1}} \ln \frac{CDM_i}{CDM_{i-1}} \quad (4)$$

3 Results

3.1 Strain and cell line development

High-producing *P. pastoris* strains and CHO cell lines expressing 3D6scFv-Fc and HSA were developed. The key data of strain and cell line development are shown in Table 1. To achieve high-level expression, amplification of the transgenes was conducted by stepwise increasing of the Zeocin concentration in case of *P. pastoris* and the MTX concentration for CHO cells according to established standard protocols.

The selection of high-producing *P. pastoris* strains was based on the determination of biomass-specific product yields in correlation to gene copy number. For 3D6scFv-Fc, the maximal product yield in shake flask screening cultures was obtained from a strain which was selected on 500 μg mL^{-1} Zeocin (Supporting information, Fig. S1A). Further increment of selection pressure resulted in decreased product yields and gene copy numbers. In contrast, maximal product yields for HSA producing strains were already reached at a Zeocin concentration of 100 μg mL^{-1} (Supporting information, Fig. S1B). Increasing antibiotic concentrations led to higher copy numbers, whereas product yields remained unchanged. Based on the desired strain properties (high product yields and high gene copy numbers), the 3D6scFv-Fc producing strain derived from 500 μg mL^{-1} Zeocin and the HSA producing strain derived from 1000 μg mL^{-1} Zeocin were chosen for bioreactor cultivation.

The established CHO clones were evaluated by monitoring the specific productivity q_P and cell growth over several passages. Increasing the MTX concentration from 0.1 to 0.4 μM improved the productivity considerably (Supporting information, Fig. S1C and D). However, a further increase of MTX concentration could not raise the specific productivity significantly. The determined gene copy numbers clearly correlated with the specific productivities. For this reason the clones at 0.4 μM MTX were selected for the final subcloning step to establish clonal high producing CHO cell lines.

3.2 Fed batch cultivation

The established high producing *P. pastoris* strains and CHO cell lines were cultivated in a fed batch process using comparable bioreactor systems (Table 1).

For *P. pastoris* cultures, the substrate feed was started after around 30 h of batch phase and was maintained for 113 h in order to reach a yeast dry mass concentration of 100 g L^{-1}. The time courses of CDM and product concentration of the 3D6scFv-Fc and HSA producing *P. pastoris* strains are shown in Fig. 1A and B.

Product concentrations were measured in the cell free culture supernatant, corrected for the volume of the biomass and recalculated to total culture volume as previously described [20]. After biomass correction, the *P. pastoris* 3D6scFv-Fc cultivation yielded 8.8 ± 0.1 mg L^{-1}. In contrast, a product concentration of 380.2 ± 32.7 mg L^{-1} could be achieved in the *P. pastoris* HSA process. The cell viability remained above 98% during the cultivations.

In case of the CHO cell lines, feeding was started after an initial batch phase of 50 h. The process was stopped after the viability dropped to 70%. CDM, cell viability and product concentration of the CHO 3D6scFv-Fc and CHO HSA cultivations are shown in Fig. 1C and D. The dura-

Table 1. Key data of *P. pastoris* and CHO system comparison

	P. pastoris	**CHO cells**
Host	SMD1168H	DUKX-B11
Promoter	GAP promoter	CMV promoter
Leader HSA	Native	Native
Leader 3D6scFv-Fc	α-factor	Human Ig heavy chain
Integration	rDNA locus	Random
Selection	Zeocin resistance	DHFR deficiency, G418 resistance
Amplification	Zeocin	Methotrexate
Bioreactor system	SR0700ODLS, DASGIP	DS0700TPSS, DASGIP
Cultivation mode	Fed batch	Fed batch
Temperature	25°C	37°C
Dissolved Oxygen	20% air saturation	30% air saturation
Stirrer speed	400–1250 rpm	80 rpm
pH	5.85	7.00
Batch medium	Synthetic, 40 g L^{-1} glycerol	DMEM/Ham's F12 based
Feed medium	Synthetic, 500 g L^{-1} glucose	DMEM/Ham's F12 based, enriched with glucose and amino acids
Feeding strategy	Glucose-limited, constant feed rate (2.38 g h^{-1})	Constant residual glucose concentration of 1–2.5 g L^{-1}, feed rate adjusted accordingly

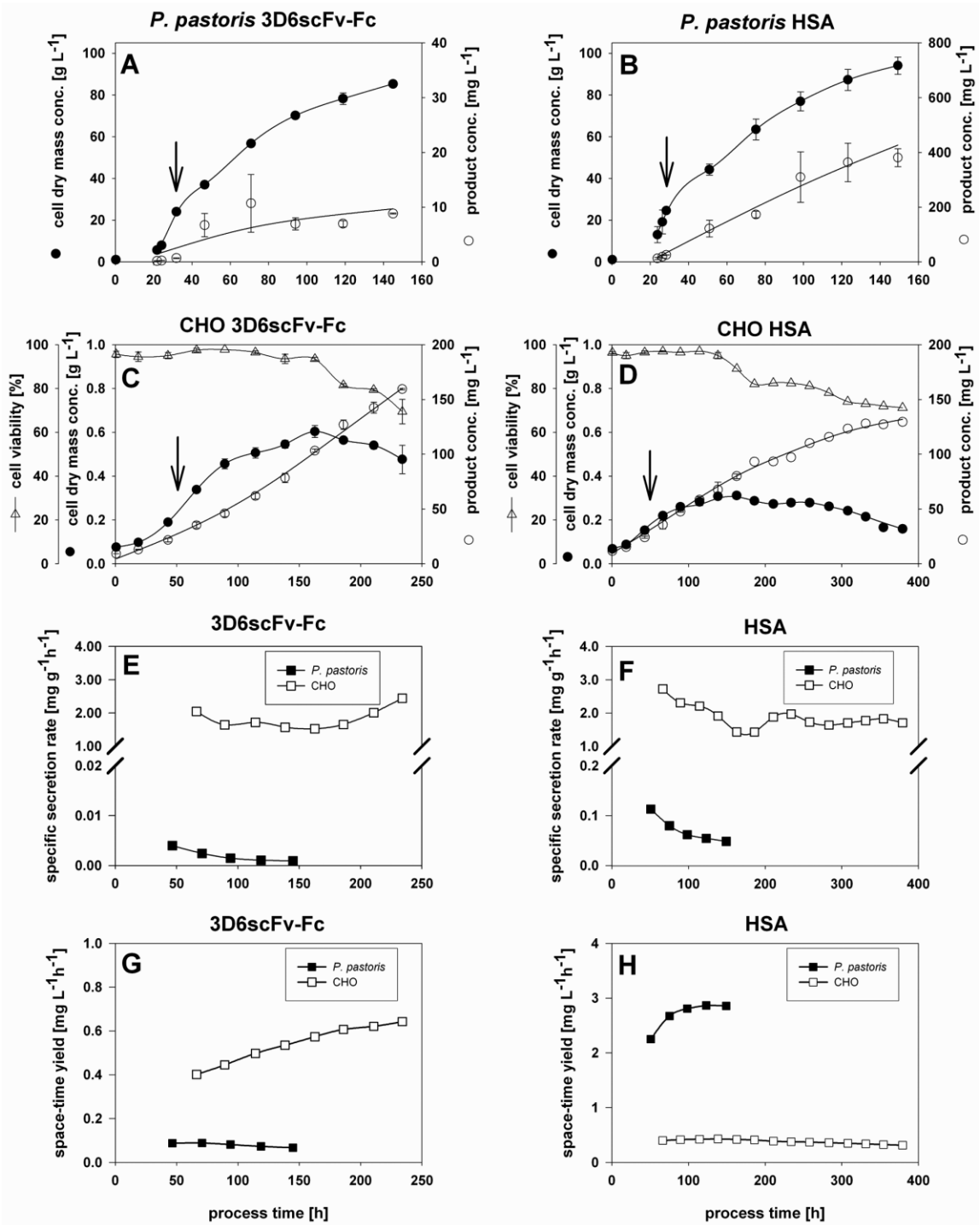


Figure 1. Time courses of fed batch cultivations. Mean cell dry mass concentration, cell viability and product concentration of the (**A**) *P. pastoris* 3D6scFv-Fc, (**B**) *P. pastoris* HSA, (**C**) CHO 3D6scFv-Fc, and (**D**) CHO HSA cultivations. Arrows mark the feed start. Data represent mean values ± standard deviation of two independent cultivations (CHO HSA: data of one cultivation after 163 h). Curves for cell dry mass and product concentration were calculated using the smoothing spline algorithm of the Matlab Curve Fitting Toolbox. Biomass-specific secretion rates of (**E**) 3D6scFv-Fc producers and (**F**) HSA producers and space-time yields of (**G**) 3D6scFv-Fc producers and (**H**) HSA producers during the feed phase. Data were calculated using smoothed product concentrations of two independent cultivations. *P. pastoris* and CHO cells were cultivated in comparable DASGIP bioreactor systems. *P. pastoris* cultures were inoculated with an optical density (OD_{600}) of 3.0 and grown at 25°C, pH 5.85, and 20% dissolved oxygen. A constant feed rate of 2.38 g h^{-1} fed batch medium (500 g L^{-1} glucose) was applied for 113 h. The CHO cell concentrations at cultivation start were adjusted to 2.5 × 10^5 cells mL^{-1} and the cultures were grown at 37°C, pH 7.0, 30% dissolved oxygen, and an agitation speed of 80 rpm. The feed rate was adjusted daily to maintain a constant residual glucose concentration of 1.0–2.5 g L^{-1}. Product concentrations were determined using ELISA and are expressed as milligram product per liter of culture broth.

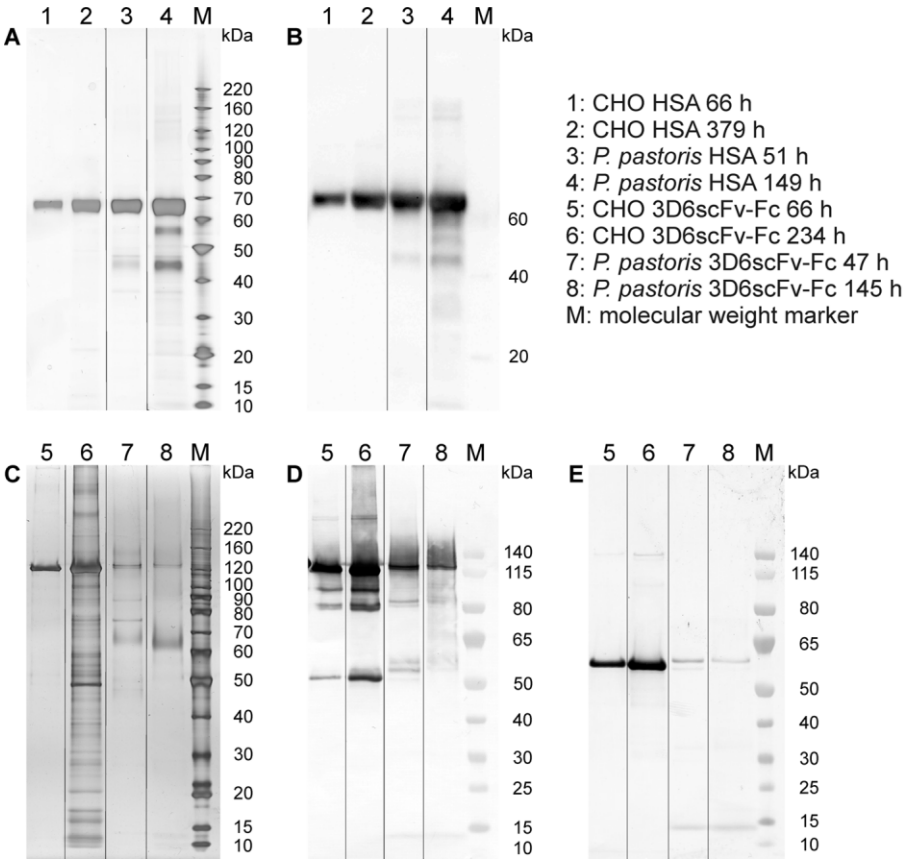


Figure 2. Fed batch culture supernatants after feed start and at the end of the process. Silver stained SDS–PAGE was used to analyze the total protein content in the culture supernatant. Western blot analysis was used to identify the recombinant proteins HSA (67 kDa) and 3D6scFv-Fc (homodimer: approx. 110 kDa), respectively. (**A**) Silver stained SDS–PAGE and (**B**) western blot under reducing conditions of the HSA producing CHO and *P. pastoris* clones. (**C**) Silver stained SDS–PAGE under non-reducing conditions, (**D**) Western blot under non-reducing, and (**E**) reducing conditions of the 3D6scFv-Fc expressing CHO and *P. pastoris* clones. Equal volumes of supernatant were loaded onto the gel. Samples of two independent cultivations were analyzed. The results of one representative replicate per clone are shown. Compared samples were always analyzed on the same gel.

tion of the CHO 3D6scFv-Fc fermentations was 234 h leading to a final product concentration of 158.9 ± 0.6 mg L^{-1}. The CHO HSA fermentation could be maintained for 379 h and yielded 129.3 mg L^{-1}.

Nearly 50-fold higher product levels were obtained for HSA in comparison to 3D6scFv-Fc with *P. pastoris*, whereas product concentrations were in a similar range for both proteins in the CHO cultivations. Notably, three- to four-fold higher final HSA concentrations were obtained in *P. pastoris* fermentations, while for 3D6scFv-Fc more than 12-fold higher product levels were reached with CHO cells.

The quality of the expressed proteins and the purity of the culture supernatant were evaluated by SDS-PAGE and western blot. The results at two points in time (after start of the feed and at the end of the process) of one representative fermentation run are shown in Fig. 2. The recombinant protein constituted the major product in the supernatant in all cultivations.

The binding affinity of the 3D6scFv-Fc antibody to the HIV-1 envelope protein UG37 gp140 was measured using bio-layer interferometry (Supporting information, Fig. S2). Very similar binding properties (k_{on}, k_{off}, and K_D values) could be observed for 3D6scFv-Fv derived from *P. pastoris* and CHO cells (Supporting information, Table S2). The determined affinity constant K_D was 3.7 nM in both cases.

3D6scFv-Fc glycosylation pattern was analyzed by mass spectrometry. CHO cell-derived 3D6scFv-Fc antibodies contained complex-type *N*-glycans which were mainly fucosylated, whereas *P. pastoris*-derived 3D6scFv-Fc antibodies had exclusively high-mannose-type *N*-glycans ranging from Man9 to Man16 (Supporting information, Fig. S3).

The quality of secreted HSA was further analyzed on the level of secondary structure. Therefore, far-UV circular dichroism spectra were recorded. The comparison to a reference albumin derived from human serum revealed no significant differences, showing that the HSA produced in *P. pastoris* and CHO cells was properly folded (Supporting information, Fig. S4).

The biomass-specific product secretion rate q_P is an essential parameter to assess the secretion performance of an expression system. Figure 1E and F show the corresponding time courses of q_P for both organisms and model proteins. In *P. pastoris*, $q_{P,mean}$ was significantly lower (40-fold) for 3D6scFv-Fc than for HSA, whereas the specific product secretion rates were very similar for both model proteins in the CHO cell lines. Furthermore, q_P clearly declined during the process in the *P. pastoris* strains. Large differences in q_P could be observed comparing both organisms. The mean specific secretion rates in the CHO cell lines were 1011-fold higher for 3D6scFv-

Table 2. Overview of the results from the fed batch cultivations

	P. pastoris 3D6scFv-Fc	CHO 3D6scFv-Fc	*P. pastoris* HSA	CHO HSA[e)]
Duration (h)	145	234	149	379
Maximal cell dry mass concentration CDM_{max} (g L^{-1})[a)]	85.3 ± 1.0	0.604 ± 0.027	94.1 ± 4.2	0.311 ± 0.009
Maximal specific growth rate μ_{max} (h^{-1})[a)]	0.147 ± 0.008	0.028 ± 0.002	0.148 ± 0.003	0.023 ± 0.002
Gene copy number	10 ± 1[c)]	18 ± 2[d)]	11 ± 2[c)]	130 ± 16[d)]
Product concentration (supernatant) (mg L^{-1})[a)]	12.3 ± 0.1	159.4 ± 0.7	551.0 ± 36.4	129.4
Product concentration (broth) (mg L^{-1})[a)]	8.8 ± 0.1	158.9 ± 0.6	380.2 ± 32.7	129.3
Mean specific secretion rate $q_{P,mean}$ (feed phase) (mg g^{-1} h^{-1})[b)]	0.0018	1.82	0.071	1.86
Maximal specific secretion rate $q_{P,max}$ (feed phase) (mg g^{-1} h^{-1})[b)]	0.0040	2.36	0.113	2.77
Space-time yield (process end) STY_{end} (mg L^{-1} h^{-1})[b)]	0.067	0.642	2.856	0.312
Maximal space-time yield STY_{max} (feed phase) (mg L^{-1} h^{-1})[b)]	0.088	0.642	2.866	0.427

a) Mean ± standard deviation ($n = 2$), two independent fed batch cultivations.
b) Calculated with smoothed product concentrations, two independent fed batch cultivations.
c) Mean ± standard deviation ($n = 4$), data from strain selection.
d) Mean ± standard deviation ($n = 2$), data from cell line selection.
e) Data of one cultivation after 163 h.

Fc and 26-fold for HSA compared to the *P. pastoris* strains (Table 2).

The STY is an important parameter to evaluate the performance of the whole process. The time courses of the STY during the feed phase for both organisms and model proteins are shown in Fig. 1G and H. In *P. pastoris* the STY at the end of the process was significantly higher for the HSA compared to the 3D6scFv-Fc expressing strain (more than 40-fold difference), whereas it was in the same range for both model proteins in the CHO cell lines (less than two-fold difference). Due to the low q_P of the 3D6scFv-Fc producing *P. pastoris* strain, the STY at the end of the process was 9.6-fold lower in *P. pastoris* in comparison with the corresponding CHO cell line. In contrast, the STY of the HSA producers was 9.2-fold higher in *P. pastoris* compared to the CHO cell line due to the shorter process time and the higher biomass density (Table 2).

All process-relevant parameters of the fed batch cultivations are summarized and compared in Table 2.

4 Discussion

In this study, we compared the ability of the two frequently applied expression systems *P. pastoris* and CHO cells for recombinant protein production. To conduct a thorough comparison, we established *P. pastoris* strains and CHO cell lines which express the same model proteins and cultivated them in fed batch processes using comparable bioreactor systems (Table 1). Both expression systems have their advantages and disadvantages considering clone development, cell growth, cell density, productivity, and product quality.

4.1 Cell line and strain development

A major advantage of the *P. pastoris* system is that stable clones can be developed within a short time. A single round of screening after transformation may be sufficient to achieve a production clone, requiring about 2 weeks. Additional rounds of amplification as described here may extend this to 2 months.

Although site-specific integration of transgenes is possible in CHO cells using recombinase-mediated cassette exchange [21, 22], random integration is still commonly used. However this necessitates extensive screening, because the chances of isolating high-producing clones depend on the number of clones that have been screened. Furthermore, the establishment of a high-producing stable CHO cell line requires steps of amplification and single cell cloning and so the whole procedure is very time-consuming and labor-intensive. Our standard procedure includes initial single-cell cloning after transfection followed by amplification, screening, and selection of the best clone. This clone will then be subjected to two additional rounds of single-cell cloning, amplification, and selection. Hence the whole clone development process from transfection to cell banking takes at least 6 months. Although high-throughput cell screening systems such as FACS-based screening have been developed, the clone development time of CHO cells cannot keep up with the very short development times of *P. pastoris* and microbial systems in general.

4.2 Cell growth and cell density

One major limitation of the CHO cell system is low specific growth rate, leading to long process times and limited cell densities. Cell growth and viable cell concentrations are additionally limited by the accumulation of tox-

ic byproducts such as ammonium [23], lactate [24] as well as elevated pCO_2 and hyperosmolality [25]. In *P. pastoris* cell density is mainly limited by the technical feasibility of biomass removal from the culture broth. Therefore one major aim of CHO process development is to achieve higher cell densities by optimizing the fed batch strategy [26], or by removing spent medium using perfusion [27]. While the low μ_{max} of CHO cells extends process duration markedly, microbial processes are commonly not limited by the cell's ability to achieve high growth rates but by technical limits of heat and oxygen transfer. It should be noted that the standard fed batch performed here for *P. pastoris* is rather long at low feed rates, leaving room for optimization.

4.3 Productivity

One means to optimize feed rates is to adapt the feed profile to the optimal trajectory of STY. We have shown before [20] that specific secretion rate and specific growth rate correlate strictly in *P. pastoris* which is also observed here. In CHO cells this correlation is rather weak (Fig. 3). Therefore different optimization strategies need to be applied to the two production platforms. In *P. pastoris*, optimization leads to initial high feed rates for rapid accumulation of biomass, followed by a phase of decreasing μ, thus allowing time for product accumulation [20]. In CHO cells, the fed batch strategy rather aims at maintaining reasonably high viable cell concentrations. This difference in feed strategy leads to large differences in process duration spanning from 50 to 150 h for *P. pastoris* up to 21 days for CHO cells [28]. The STY is a measure for the product output per bioreactor volume and time, and thus illustrates the respective capital costs per unit of product. STY is reverse proportional to process duration, so that in other words the capital costs per unit product increase proportionally with the time needed to achieve a defined amount of product. Especially for the biopharmaceutical industry capital costs for production plants are a major factor of total production costs [29], so that maximizing STY is a valid optimization strategy.

Our results illustrate clearly that specific productivity (mean values during feed phase) in *P. pastoris* depends strongly on the product, being 40-fold lower for the antibody fragment compared to HSA. Low q_P of the antibody fragment cannot be explained by protein size and number of disulfide bonds as HSA is a large protein of 67 kDa composed of three domains with a total of 17 intramolecular disulfide bonds and one free cysteine, compared to about 110 kDa and 10 disulfide bonds of the dimeric antibody fragment. It is more likely that folding efficiency and thermodynamic stability account for differences in q_P as described for different variants of scFv fragments [30]. Notably this does not account for CHO cells as q_P is the same for both proteins in this expression system. It may be speculated that CHO cells are specifically better suited to produce antibodies or their fragments while other proteins like HSA are well produced also in lower eukaryotic expression systems.

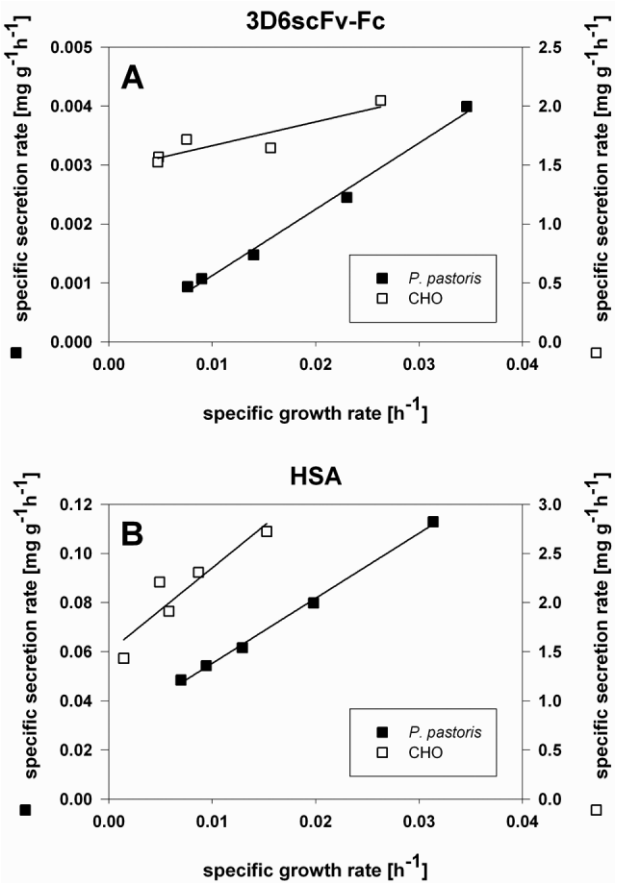


Figure 3. Relationship between biomass-specific secretion rate q_P and specific growth rate μ during the feed phase of fed batch cultivations. (**A**) 3D6scFv-Fc and (**B**) HSA producing *P. pastoris* and CHO clones. *P. pastoris* and CHO cells were cultivated as described for Fig. 1. Product concentrations were determined using ELISA. Biomass-specific secretion rates were calculated using smoothed product concentrations (smoothing spline algorithm of the Matlab Curve Fitting Toolbox) of two independent cultivations. Specific growth rates represent mean values of two independent cultivations.

4.4 Product quality

Downstream processing is the major cost factor in a biopharmaceutical production process accounting for 50–80% of the total manufacturing costs [31]. Hence, it is very important that an expression system produces the desired protein at high relative purity to facilitate purification. Both production systems in this study are able to deliver very pure supernatants that contain low amounts of host cell proteins. Certainly, the produced proteins should be correctly folded and fully assembled to ensure biological activity. In the CHO processes we did not observe any product degradation. In case of *P. pastoris*

very little amounts of degraded HSA could be detected. However, most of the HSA had the correct size and was properly folded. The 3D6scFv-Fc antibody was predominantly assembled as a dimer of the correct size in both expression systems. Not surprisingly, differences could be observed in the *N*-glycosylation pattern of the 3D6scFv-Fc antibody. The *P. pastoris*-derived antibodies were partly unglycosylated or contained glycans of the high-mannose-type. In contrast, the Fc glycans present in antibodies produced in CHO cells had the complex-type biantennary structures which are typically found in human IgGs [32]. As expected, the differences in IgG-Fc glycosylation did not impair binding to the epitope. However, it is well known that Fcγ receptor and complement mediated effector functions (antibody-dependent cellular cytotoxicity and complement-dependent cytotoxicity) are strongly affected by the IgG glycoform [33]. Hence, glycoengineered *P. pastoris* strains [34] would be required to produce fully functional IgGs or scFv-Fc antibodies.

4.5 Concluding remarks

In this study, we could demonstrate that in contrast to CHO cells the secretion capacity of *P. pastoris* is highly dependent on the complexity of the produced recombinant protein. In *P. pastoris* the specific product secretion rates are considerably lower compared to CHO cells. However, the shorter process times and the higher cell densities of *P. pastoris* processes can compensate for the lower secretion potential. Whether *P. pastoris* or CHO cells are the more adequate expression system strongly depends on the particular protein. Considering the two model proteins used in this study, we concluded that antibodies and antibody fragments may be produced more economically in CHO cells. On the contrary, *P. pastoris* seems to be the better choice for the production of less complex proteins such as HSA.

The results of this study clearly indicate that the protein secretion machinery is much more efficient in mammalian cells than in yeast. The mechanisms that control and limit recombinant protein expression and secretion in yeast as well as in mammalian cells are still poorly understood. Hence, a comprehensive comparison of these expression systems on a physiological level using various omics technologies might be very useful to identify the bottlenecks of heterologous protein secretion in lower eukaryotes.

This work has been supported by the Federal Ministry of Economy, Family and Youth (BMWFJ), the Federal Ministry of Traffic, Innovation and Technology (bmvit), the Styrian Business Promotion Agency SFG, the Standortagentur Tirol and ZIT – Technology Agency of the City of Vienna through the COMET-Funding Program managed by the Austrian Research Promotion Agency FFG. This work has also been funded by the PhD program "BioToP – Biomolecular Technology of Proteins" (Austrian Science Fund, FWF Project W1224).

We thank Stefan Bauernfried for his excellent assistance with the fermentations of the CHO cell lines and are grateful to Stefan Ertl and Gabriele Wilt for development of P. pastoris strains and excellent analytical support respectively, and to Sabine Necina for help with the ÄKTA chromatography system.

The authors declare no financial or commercial conflict of interest.

5 References

[1] Ferrer-Miralles, N., Domingo-Espín, J., Corchero, J. L., Vázquez, E., Villaverde, A., Microbial factories for recombinant pharmaceuticals. *Microb. Cell Fact.* 2009, *8*, 17.

[2] Corchero, J. L., Gasser, B., Resina, D., Smith, W. et al., Unconventional microbial systems for the cost-efficient production of high-quality protein therapeutics. *Biotechnol Adv* 2013, *31*, 140–153.

[3] Hacker, D. L., De Jesus, M., Wurm, F. M., 25 years of recombinant proteins from reactor-grown cells – where do we go from here? *Biotechnol. Adv.* 2009, *27*, 1023–1027.

[4] Lai, T., Yang, Y., Ng, S., Advances in mammalian cell line development technologies for recombinant protein production. *Pharmaceuticals* 2013, *6*, 579–603.

[5] Waegeman, H., Soetaert, W., Increasing recombinant protein production in *Escherichia coli* through metabolic and genetic engineering. *J. Ind. Microbiol. Biotechnol.* 2011, *38*, 1891–1910.

[6] Mattanovich, D., Branduardi, P., Dato, L., Gasser, B. et al., Recombinant protein production in yeasts. *Methods Mol. Biol.* 2012, *824*, 329–358.

[7] Dasgupta, S., Navarrete, A. M., Bayry, J., Delignat, S. et al., A role for exposed mannosylations in presentation of human therapeutic self-proteins to CD4+ T lymphocytes. *Proc. Natl. Acad. Sci. USA* 2007, *104*, 8965–8970.

[8] Sola, R. J., Griebenow, K., Effects of glycosylation on the stability of protein pharmaceuticals. *J. Pharm. Sci.* 2009, *98*, 1223–1245.

[9] Hamilton, S. R., Davidson, R. C., Sethuraman, N., Nett, J. H. et al., Humanization of yeast to produce complex terminally sialylated glycoproteins. *Science* 2006, *313*, 1441–1443.

[10] Potgieter, T. I., Cukan, M., Drummond, J. E., Houston-Cummings, N. R. et al., Production of monoclonal antibodies by glycoengineered *Pichia pastoris*. *J. Biotechnol.* 2009, *139*, 318–325.

[11] Kobayashi, K., Kuwae, S., Ohya, T., Ohda, T. et al., High-level expression of recombinant human serum albumin from the methylotrophic yeast *Pichia pastoris* with minimal protease production and activation. *J. Biosci. Bioeng.* 2000, *89*, 55–61.

[12] Felgenhauer, M., Kohl, J., Rüker, F., Nucleotide sequences of the cDNAs encoding the V-regions of H- and L-chains of a human monoclonal antibody specific to HIV-1-gp41. *Nucleic Acids Res.* 1990, *18*, 4927.

[13] Stadlmayr, G., Mecklenbräuker, A., Rothmüller, M., Maurer, M. et al., Identification and characterisation of novel *Pichia pastoris* promoters for heterologous protein production. *J. Biotechnol.* 2010, *150*, 519–529.

[14] Marx, H., Mecklenbräuker, A., Gasser, B., Sauer, M., Mattanovich, D., Directed gene copy number amplification in *Pichia pastoris* by vector integration into the ribosomal DNA locus. *FEMS Yeast Res.* 2009, *9*, 1260–1270.

[15] Urlaub, G., Chasin, L. A., Isolation of Chinese-hamster cell mutants deficient in dihydrofolate-reductase activity. *Proc. Natl. Acad. Sci. Biol.* 1980, *77*, 4216–4220.
[16] Reisinger, H., Steinfellner, W., Katinger, H., Kunert, R., Serum-free transfection of CHO cells with chemically defined transfection systems and investigation of their potential for transient and stable transfection. *Cytotechnology* 2009, *60*, 115–123.
[17] Prielhofer, R., Maurer, M., Klein, J., Wenger, J. et al., Induction without methanol: Novel regulated promoters enable high-level expression in *Pichia pastoris*. *Microb. Cell Fact.* 2013, *12*, 5.
[18] Hohenblum, H., Borth, N., Mattanovich, D., Assessing viability and cell-associated product of recombinant protein producing *Pichia pastoris* with flow cytometry. *J. Biotechnol.* 2003, *102*, 281–290.
[19] Heukeshoven, J., Dernick, R., Simplified method for silver staining of proteins in polyacrylamide gels and the mechanism of silver staining. *Electrophoresis* 1985, *6*, 103–112.
[20] Maurer, M., Kuhleitner, M., Gasser, B., Mattanovich, D., Versatile modeling and optimization of fed batch processes for the production of secreted heterologous proteins with *Pichia pastoris*. *Microb. Cell Fact.* 2006, *5*, 37.
[21] Kim, J. Y., Kim, Y. G., Lee, G. M., CHO cells in biotechnology for production of recombinant proteins: Current state and further potential. *Appl. Microbiol. Biotechnol.* 2012, *93*, 917–930.
[22] Wirth, D., Gama-Norton, L., Riemer, P., Sandhu, U. et al., Road to precision: Recombinase-based targeting technologies for genome engineering. *Curr. Opin. Biotechnol.* 2007, *18*, 411–419.
[23] Xing, Z., Li, Z., Chow, V., Lee, S. S., Identifying inhibitory threshold values of repressing metabolites in CHO cell culture using multivariate analysis methods. *Biotechnol. Prog.* 2008, *24*, 675–683.
[24] Lao, M. S., Toth, D., Effects of ammonium and lactate on growth and metabolism of a recombinant Chinese hamster ovary cell culture. *Biotechnol. Prog.* 1997, *13*, 688–691.
[25] Zhu, M. M., Goyal, A., Rank, D. L., Gupta, S. K. et al., Effects of elevated pCO2 and osmolality on growth of CHO cells and production of antibody-fusion protein B1: A case study. *Biotechnol. Prog.* 2005, *21*, 70–77.
[26] Huang, Y. M., Hu, W., Rustandi, E., Chang, K. et al., Maximizing productivity of CHO cell-based fed-batch culture using chemically defined media conditions and typical manufacturing equipment. *Biotechnol. Prog.* 2010, *26*, 1400–1410.
[27] Clincke, M. F., Molleryd, C., Samani, P. K., Lindskog, E. et al., Very high density of Chinese hamster ovary cells in perfusion by alternating tangential flow or tangential flow filtration in WAVE bioreactor-part II: Applications for antibody production and cryopreservation. *Biotechnol. Prog.* 2013, *29*, 768–777.
[28] De Jesus, M., Wurm, F. M., Manufacturing recombinant proteins in kg-ton quantities using animal cells in bioreactors. *Eur. J. Pharm. Biopharm.* 2011, *78*, 184–188.
[29] Werner, R. G., Economic aspects of commercial manufacture of biopharmaceuticals. *J. Biotechnol.* 2004, *113*, 171–182.
[30] Whyteside, G., Alcocer, M. J., Kumita, J. R., Dobson, C. M. et al., Native-state stability determines the extent of degradation relative to secretion of protein variants from *Pichia pastoris*. *PLoS ONE* 2011, *6*, e22692.
[31] Roque, A. C. A., Lowe, C. R., Taipa, M. A., Antibodies and genetically engineered related molecules: Production and purification. *Biotechnol. Progr.* 2004, *20*, 639–654.
[32] Zauner, G., Selman, M. H. J., Bondt, A., Rombouts, Y. et al., Glycoproteomic analysis of antibodies. *Mol. Cell Proteomics* 2013, *12*, 856–865.
[33] Jefferis, R., Recombinant antibody therapeutics: The impact of glycosylation on mechanisms of action. *Trends Pharmacol. Sci.* 2009, *30*, 356–362.
[34] Ye, J. X., Ly, J., Watts, K., Hsu, A. et al., Optimization of a glycoengineered *Pichia pastoris* cultivation process for commercial antibody production. *Biotechnol. Progr.* 2011, *27*, 1744–1750.

Biotechnology Journal

Supporting Information for DOI 10.1002/biot.201300305

Pichia pastoris secretes recombinant proteins less efficiently than Chinese hamster ovary cells but allows higher space-time yields for less complex proteins

Andreas Maccani, Nils Landes, Gerhard Stadlmayr, Daniel Maresch, Christian Leitner, Michael Maurer, Brigitte Gasser, Wolfgang Ernst, Renate Kunert, Diethard Mattanovich

Supporting information: Strain and cell line selection

Supporting materials and methods

Transgene copy number determination in *P. pastoris*

Isolation of genomic DNA and determination of gene copy numbers was done as described previously [1]. Briefly, genomic DNA was isolated using the DNeasy Blood & Tissue Kit (Qiagen, Germany) according to the supplier's supplementary protocol. Concentration of isolated DNA was measured with a spectrophotometer at 260nm (Nanodrop). Gene copy numbers were determined with quantitative real-time PCR using the SensiMix SYBR Kit (Bioline, UK) and a Rotor-Gene PCR cycler (Qiagen, Germany). Each reaction was performed in quadruplicate. Raw data were analyzed relative to a reference sample with known gene copy number by using the comparative quantitation method of the Rotor-Gene system software (Qiagen, Germany). The protein disulfide isomerase (PDI) gene was used as normalizer. Primers used in this study are given in Table S1.

Transgene copy number determination in CHO cells

Cells at different stages of gene amplification were cultivated in 250 mL Erlenmeyer flasks at 37 °C, 7 % CO_2, 90 % humidity and 120 rpm in an incubator shaker. Samples for gene copy number (GCN) determination were taken in the exponential growth phase of two consecutive passages. 5×10^6 cells were washed with Dulbecco's PBS and genomic DNA was isolated using the QIAamp DNA Blood Mini Kit (Qiagen, Germany) according to the manufacturer's instructions including the RNase A digestion step in order to obtain RNA-free DNA samples. The genomic DNA was then denatured at 99 °C for 10 min and stored at 4 °C. DNA concentration and purity (A_{260}/A_{280}) was determined using the NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific, MA, USA). The GCN was determined by quantitative real-time PCR (qPCR) using a SYBR Green I assay and the MiniOpticon real-time PCR detection system (Bio-Rad, CA, USA). One set of primers specific to the CMV immediate-early enhancer/promoter was employed for all cell lines (Table S1). The amplicon size was 131 bp. The linearized pCI-neo expression vector was used as standard for absolute quantification of the copy number. Standard dilutions containing 1×10^7 to 1×10^2 copies in a 1:10 dilution series were analyzed in triplicates for each qPCR run. Genomic DNA samples were analyzed in triplicates using 10 ng as a template. The unreplicated nuclear DNA amount of CHO cells was estimated to be 5.4 pg [2] and so the analyzed 10 ng represent the nuclear DNA amount of 1850 cells. The 20 µL reaction mix consisted of the SensiMix SYBR No-ROX Kit (Bioline, UK), 250 nM of each oligonucleotide primer and the appropriate amount of template DNA. qPCR runs were performed by 10 min initial polymerase activation and DNA denaturation at 95 °C followed by 40 cycles of amplification (15 s at 95 °C, 15 s at 66 °C and 15 s at 72 °C). Fluorescence was recorded after each amplification cycle. The specificity of the reactions was verified by analyzing the melting curve immediately after the last amplification cycle. The results were evaluated with the CFX Manager software 2.1 (Bio-Rad, CA, USA). Quantification cycle (C_q) values were determined using the 'Regression'-mode. PCR efficiency was calculated from the slope of the standard curve and was within 96.0 and 97.5 % for all qPCR runs.

Supporting references

[1] Marx, H., Mecklenbräuker, A., Gasser, B., Sauer, M., Mattanovich, D., Directed gene copy number amplification in *Pichia pastoris* by vector integration into the ribosomal DNA locus. *FEMS Yeast Res* 2009, *9*, 1260-1270.

[2] Maccani, A., Ernst, W., Grabherr, R., Whole Genome Sequencing Improves Estimation of Nuclear DNA Content of Chinese Hamster Ovary Cells. *Cytom Part A* 2013, *83*, 893-895.

Supporting tables and figures

Table S1 Sequences of oligonucleotides used for determination of gene copy numbers by quantitative real-time PCR

Primer	Organism	Sequence (5´–3´)
3D6 fw	*P. pastoris*	TGC TGT TGA ATG GGA GTC TAA
3D6 rev	*P. pastoris*	GCA TAA CGG AAC AGG AGA AAA C
HSA fw	*P. pastoris*	AGA GAG ACA GAT CAA GAA GCA GAC
HSA rev	*P. pastoris*	CCA ACT TCT TAC CCT CTT CAG CG
PDI fw	*P. pastoris*	GGA AAG GCC CAC GAT GAA GTT GTC
PDI rev	*P. pastoris*	GCA TCC TCA TCA TTG GCG TAA AGA GTA G
CMV fw	CHO cells	ACC AAT GGG CGT GGA TAG CG
CMV rev	CHO cells	GGG CGA TCG CAG TTG TTA CG

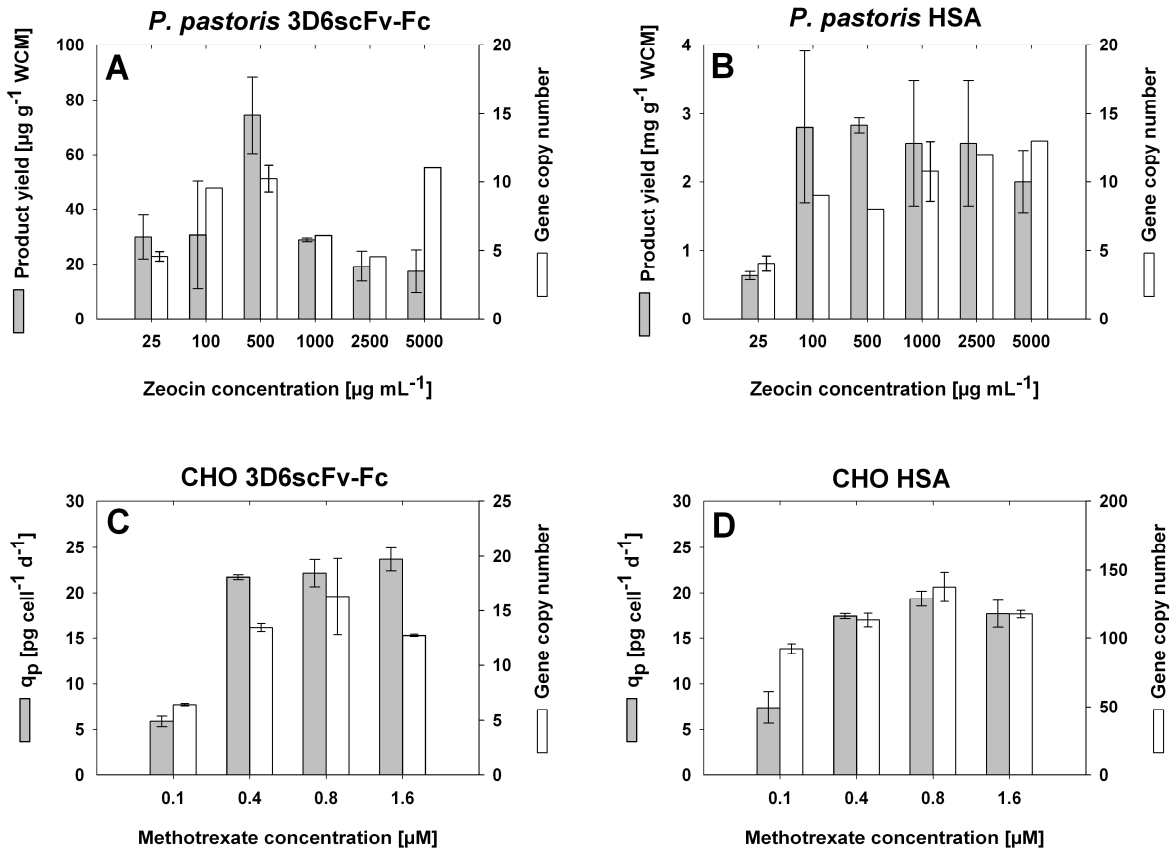


Figure S1 Strain and cell line selection. Selection of *P. pastoris* strains was based on biomass-specific (wet cell mass) product yields and gene copy numbers (A and B). *P. pastoris* samples were taken after 48 h of shake flask cultivation. Product yields for *P. pastoris* represent mean values of two independent screenings. Gene copy numbers for 3D6scFv-Fc producing strains derived from 25 and 500 µg mL^{-1} Zeocin as well as for HSA producing strains derived from 25 and 1000 µg mL^{-1} Zeocin represent mean values of one screening and three independent bioreactor cultivations and confirm the reproducibility of the experimental procedure and the genetic stability of the generated yeast strains. For all other *P. pastoris* strains gene copy number determination was based on one biological sample from a screening culture. Shown error bars represent standard deviations of two or four replicates as described above. For CHO cell lines specific product secretion rates (q_P) and gene copy numbers were the primary selection criteria (C and D). Data for CHO cells were obtained form shake flask cultures and represent mean values ± standard deviation of three (q_P) and two (gene copy number) consecutive passages. Product concentrations for yield and q_P calculation were measured using ELISA. Gene copy numbers were determined by quantitative real-time PCR and represent absolute values per unreplicated genome.

Supporting information: 3D6scFv-Fc binding affinity

Supporting materials and methods

Binding affinity determination

Binding affinity of 3D6scFv-Fc to the HIV-1 envelope protein UG37 gp140 (Polymun Scientific, Austria) was determined using bio-layer interferometry (BLI). The Octet QK instrument (Pall ForteBio, CA, USA) with Protein A biosensors was used to perform the kinetic assays. The 3D6scFv-Fc antibodies were loaded onto the biosensors and the kinetic measurements were performed at serial dilutions of UG37 gp140 (2 µM, 1 µM, 500 nM and 250 nM). The assays were conducted at 30 °C and 1000 rpm agitation in Pall ForteBio's Kinetics Buffer. All data were processed by reference subtraction, y-axis alignment, inter-step correction and Savitzky-Golay filtering. Binding kinetics were calculated using the 1:1 binding model and global fit analysis provided by the Octet data analysis software 6.4 (Pall ForteBio, CA, USA).

Supporting tables and figures

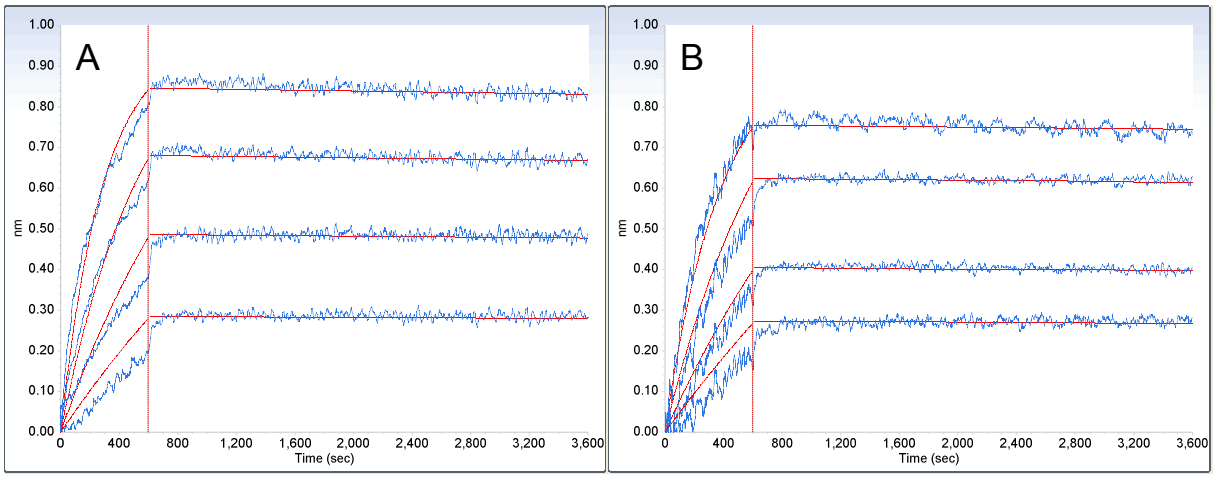


Figure S2 Affinity measurement of 3D6scFv-Fc to HIV-1 envelope protein UG37 gp140. Bio-layer interferometry sensorgrams showing association and dissociation of UG37 gp140 to 3D6scFv-Fc derived from (A) *P. pastoris* and (B) CHO cells. UG37 gp140 was applied at concentrations of 2 µM, 1 µM, 500 nM and 250 nM (blue curves). Data was fitted using the 1:1 binding model (red curves).

Table S2 Kinetics for binding of 3D6scFv-Fc to the HIV-1 envelope protein UG37 gp140

	P. pastoris 3D6scFv-Fc	CHO 3D6scFv-Fc
k_{on} [10^3 M^{-1} s^{-1}]	1.97 ± 0.25	1.32 ± 0.01
k_{off} [10^{-6} s^{-1}]	7.21 ± 0.84	4.90 ± 0.19
K_D [nM]	3.66 ± 0.03	3.71 ± 0.13

Mean ± standard deviation (n = 2), samples of two independent fed batch cultivations

Supporting information: 3D6scFv-Fc glycosylation

Supporting materials and methods

HPLC-ESI-MS analysis of glycopeptides

Supernatants from fed batch processes were used for determination of the N-glycosylation pattern of the 3D6scFv-Fc antibody. Antibodies were subjected to standard SDS-PAGE under non-reducing conditions. Coomassie stained 3D6scFv-Fc homodimers were excised, S-alkylated and digested with trypsin as described elsewhere [3]. Prior to SDS-PAGE separation, *P. pastoris* supernatant was concentrated 10-fold by acetone precipitation to provide a sufficient amount of antibody for subsequent analysis steps. Glycopeptide analysis was performed on a high-performance liquid chromatography electrospray ionization mass spectrometry (HPLC-ESI-MS) system consisting of a Dionex UltiMate 3000 LC system (Thermo Fisher Scientific, MA, USA) and a Waters Q-TOF Ultima MS system (Waters, MA, USA) with standard ESI-source. Glycopeptides were separated on a BioBasic C18 column (150 mm x 0.32 mm, 5 µm, Thermo Fisher Scientific, MA, USA). A gradient from 1 % to 80 % acetonitrile (ACN) was developed over 60 min at a flow rate of 6 µL min^{-1} (solvent A consisted of 0.3 % formic acid buffered to pH 3.0 with ammonia and solvent B was 100 % ACN). Positive ions in the range from *m*/*z* 400 to 1800 were measured. Capillary voltage was 3.2 kV and time of flight voltage 9.1 kV, source temperature was 100 °C and desolvation temperature 200 °C. Data were evaluated using MassLynx 4.0 software.

Supporting references

[3] Stadlmann, J., Pabst, M., Kolarich, D., Kunert, R., Altmann, F., Analysis of immunoglobulin glycosylation by LC-ESI-MS of glycopeptides and oligosaccharides. *Proteomics* 2008, *8*, 2858-2871.

Supporting tables and figures

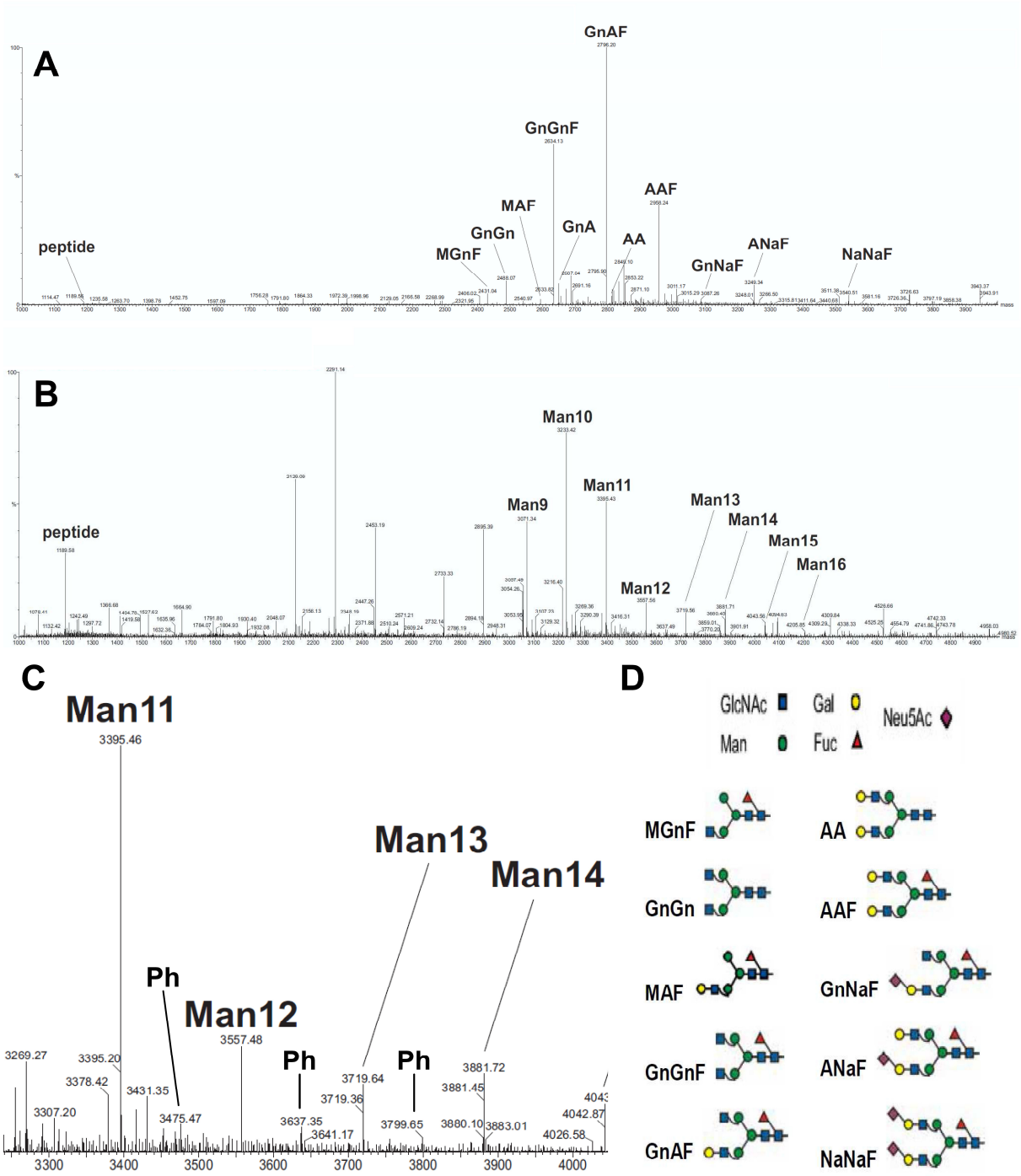


Figure S3 N-glycosylation patterns of secreted 3D6scFv-Fc. The deconvoluted glycopeptide spectra of (A) CHO cell and (B) *P. pastoris* produced 3D6scFv-Fc was determined by high-performance liquid chromatography electrospray ionization mass spectrometry (HPLC-ESI-MS). CHO cell derived 3D6scFv-Fc carries several complex-type N-glycans (mostly fucosylated), whereas *P. pastoris* derived 3D6scFv-Fc carries exclusively high-mannose-type N-glycans, ranging from Man9 to Man16. In contrast to CHO, a notable amount of non-glycosylated product (indicated by the peptide peak in (A) and (B)) was found for *P. pastoris*. Panel (C) highlights the presence of phosphorylated forms of the oligo-mannosidic N-glycans Man11, Man12 and Man13 on *P. pastoris* produced 3D6scFv-Fc, albeit at very low levels when compared to the abundance of the corresponding non-phosphorylated mannose structures. The relevant part of the deconvoluted glycopeptide spectra is depicted. Peaks representing phosphorylated mannose forms are marked with Ph (+ 80 Da distance to the respective non-phosphorylated structure). Y-axes of glycopeptide spectra show relative intensity in % and X-axes represent mass of $[M+H]^+$ ion. Panel (D) summarizes the structures of the complex-type N-glycans found in this study. N-glycan nomenclature is based on the Proglycan abbreviation system (www.proglycan.com). Abbreviations (in brackets: Proglycan single letter code): GlcNAc (Gn) – *N*-acetylglucosamine; Gal (A) – galactose; Man (M) – mannose; Fuc (F) – fucose; Neu5Ac (Na) – *N*-acetylneuraminic acid.

Supporting information: HSA structure

Supporting materials and methods

HSA purification and Electronic Circular Dichroism analysis

To investigate the secondary structure of secreted HSA, supernatants from fed batch processes were purified and applied to Electronic Circular Dichroism (ECD) analysis. Purification of supernatants was done using a 1 mL HiTrap Blue HP column in combination with an ÄKTA chromatography system (GE Healthcare, UK) following the manufacturer`s protocol. For each organism 40 mL of filtered protein solution, containing a total amount of 3.6 mg HSA, were loaded onto the column by applying a flow rate of 0.7 mL h^{-1}. Prior to loading, protein solutions were filtered (Millex GP, 0.22 µm, EMD Millipore, MA, USA). 50 mM KH_2PO_4 solution was used as binding buffer (pH 7.0). Elution buffer contained 50 mM KH_2PO_4 and 1.5 M KCl (pH 7.0). HSA was eluted from the column with 5 mL elution buffer at a flow rate of 0.7 mL h^{-1}. Desalting of eluates was achieved by using PD-10 Desalting columns (GE Healthcare, UK) following the gravity protocol according to the manufacturer's instructions. 50 mM KH_2PO_4 solution (pH 7.0) was used as elution buffer. Amicon Ultra-0.5 Centrifugal Filter devices (EMD Millipore, MA, USA) were used for further concentration of the protein solutions according to the supplier's manual. ECD spectra were measured on a Chirascan spectropolarimeter (Applied Photophysics, UK) that allowed simultaneous UV-Vis and ECD monitoring. The system was equipped with a Peltier temperature control unit and flushed with nitrogen at a flow rate of 3 L min^{-1}. Far-UV spectra (190–280 nm) were measured in a 1 mm path-length quartz cuvette. The following instrument settings were used: spectral bandwidth: 1 nm; step size: 1 nm; scan time: 15 min. The protein concentration was 0.30 mg mL^{-1}. All ECD measurements were performed in binding buffer at 25 °C. Each spectrum was automatically corrected with the base line to remove birefringence of the cell.

Supporting tables and figures

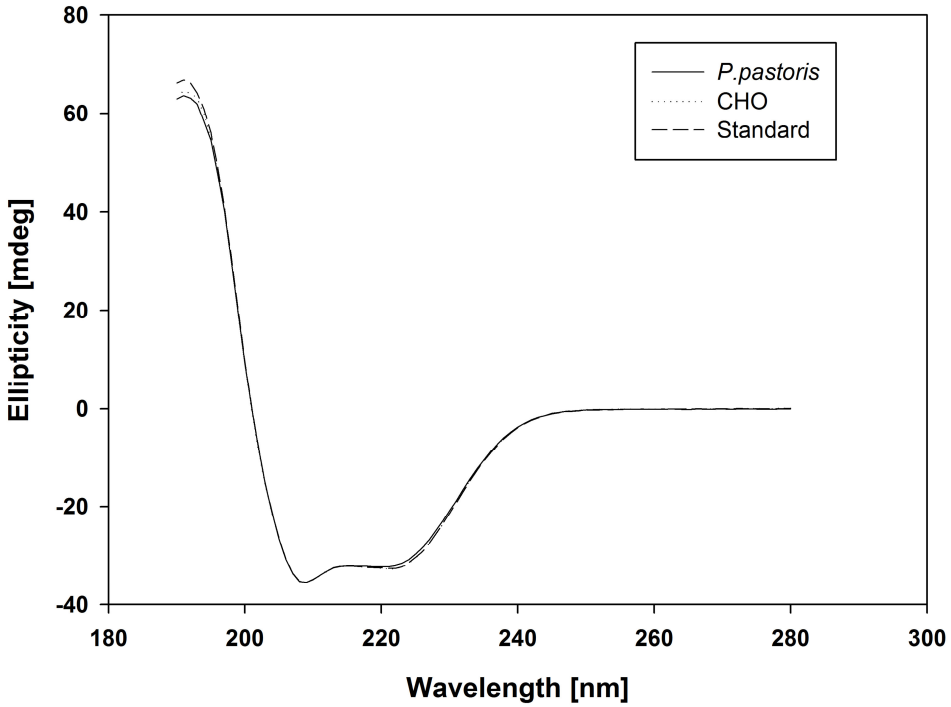


Figure S4 Electronic circular dichroism based characterization of secreted human serum albumin (HSA). Far-UV spectra (190–280 nm) of HSA produced in *P. pastoris* and CHO cells were compared to a human serum derived HSA standard (A8763, Sigma-Aldrich, MO, USA).

3.3 Publication C: MicroRNA profiling of recombinant CHO cells

Identification of microRNAs specific for high producer CHO cell lines using steady-state cultivation

Andreas Maccani, Matthias Hackl, Christian Leitner, Willibald Steinfellner, Alexandra B. Graf, Nadine E. Tatto, Michael Karbiener, Marcel Scheideler, Johannes Grillari, Diethard Mattanovich, Renate Kunert, Nicole Borth, Reingard Grabherr and Wolfgang Ernst

Accepted for publication in Applied Microbiology and Biotechnology

Accepted 24 June 2014

DOI: 10.1007/s00253-014-5911-4

GENOMICS, TRANSCRIPTOMICS, PROTEOMICS

Identification of microRNAs specific for high producer CHO cell lines using steady-state cultivation

Andreas Maccani[1,2], Matthias Hackl[2], Christian Leitner[3], Willibald Steinfellner[2], Alexandra B. Graf[1,4], Nadine E. Tatto[1], Michael Karbiener[5], Marcel Scheideler[5], Johannes Grillari[2], Diethard Mattanovich[1,2], Renate Kunert[1,2], Nicole Borth[1,2], Reingard Grabherr[2], Wolfgang Ernst[1,2]

[1] Austrian Centre of Industrial Biotechnology (ACIB GmbH), Muthgasse 11, 1190 Vienna, Austria

[2] Department of Biotechnology, VIBT – BOKU – University of Natural Resources and Life Sciences Vienna, Muthgasse 18, 1190 Vienna, Austria

[3] Department of Food Science and Technology, VIBT – BOKU – University of Natural Resources and Life Sciences Vienna, Muthgasse 18, 1190 Vienna, Austria

[4] School of Bioengineering, University of Applied Sciences FH-Campus Vienna, Muthgasse 56, 1190 Vienna, Austria

[5] Institute for Genomics and Bioinformatics, Graz University of Technology, Petersgasse 14, 8010 Graz, Austria

Correspondence: Andreas Maccani, ACIB GmbH, Muthgasse 11, 1190 Vienna, Austria
Email: andreas.maccani@acib.at
Tel: +43 1 47654 6908
Fax: +43 1 47654 6675

Abstract

MicroRNAs are short non-coding RNAs that play an important role in the regulation of gene expression. Hence, microRNAs are considered as potential targets for engineering of Chinese hamster ovary (CHO) cells to improve recombinant protein production. Here, we analyzed and compared the microRNA expression patterns of high, low and non-producing recombinant CHO cell lines expressing two structurally different model proteins in order to identify microRNAs that are involved in heterologous protein synthesis and secretion and thus might be promising targets for cell engineering to increase productivity. To generate reproducible and comparable data, the cells were cultivated in a bioreactor under steady-state conditions. Global microRNA expression analysis showed that mature microRNAs were predominantly up-regulated in the producing cell lines compared to the non-producer. Several microRNAs were significantly differentially expressed between high and low producers, but none of them commonly for both model proteins. The identification of target mRNAs is essential to understand the biological function of microRNAs. Therefore, we negatively correlated microRNA and global mRNA expression data and combined them with computationally predicted and experimentally validated targets. However, statistical analysis of the identified microRNA-mRNA interactions indicated a considerable false positive rate. Our results and the comparison to published data suggest that the reaction of CHO cells to heterologous protein expression is strongly product and/or clone-specific. In addition, this study highlights the urgent need for reliable CHO-specific microRNA target prediction tools and experimentally validated target databases in order to facilitate functional analysis of high-throughput microRNA expression data in CHO cells.

Keywords: Chinese hamster ovary cells, productivity, chemostat, microarray, miRNA expression profiling, miRNA target identification

Introduction

Chinese hamster ovary (CHO) cells are the most frequently applied expression system for the production of therapeutic proteins, mainly because of their ability to grow in suspension and to secrete complex recombinant proteins that are correctly processed. CHO cells allow proper protein folding and post-translational modifications such as human-like glycosylation which might be required for biological efficacy. Mammalian cells have originally been considered as the least effective production systems. But due to the advances in cell culture technology over the last three decades, product titers of 1 to 5 g L^{-1} are typically reached in industry today (Hacker et al. 2009). This was mainly achieved by vector, media and process optimization, but also cell engineering has been applied to improve the productivity of recombinant CHO cells. A multitude of strategies to engineer apoptosis resistance, cell proliferation, product secretion or cell metabolism have been described and are comprehensively reviewed elsewhere (Kim et al. 2012). These approaches often include the stable overexpression of one or more genes, which constitutes an additional burden to the translational machinery of the cell. To circumvent this drawback, microRNAs (miRNAs) have been considered as potential targets for cell engineering (Müller et al. 2008). MicroRNAs are short (~22 nt) endogenous RNAs that play an important role in the regulation of gene expression (Bartel 2004). They are predominantly transcribed by RNA Polymerase II, processed by Drosha and exported to the cytoplasm where the ~70 nt precursor hairpin (pre-miRNA) is cleaved by Dicer resulting in a miRNA/miRNA* duplex. One strand (mature miRNA) associates with Argonaute (AGO) proteins and forms a miRNA-induced silencing complex (miRISC) which recognizes the target mRNAs predominantly by binding to the 3'-UTR through imperfect base-pairing (Krol et al. 2010). MicroRNAs act in a post-transcriptional manner by decreasing translational efficiency and/or transcript levels. A single miRNA can repress hundred different mRNAs and thereby regulates entire gene networks (Hobert 2008). They play crucial roles in a wide range of biological processes including development, proliferation, differentiation, apoptosis and metabolism (Bartel 2004; He and Hannon 2004). In mammalian cells, miRNAs are predicted to regulate or fine-tune gene expression of ~50% of all protein-coding genes (Krol et al. 2010). In CHO cells it has already been shown that miRNAs can be utilized to improve growth (Jadhav et al. 2012), apoptosis resistance (Druz et al. 2013) and specific productivity (Barron et al. 2011; Jadhav et al. 2014; Strotbek et al. 2013).

In this study, we analyzed and compared the miRNA expression pattern of high, low and non-producing recombinant CHO cell lines to identify miRNA targets that are involved in recombinant protein synthesis and secretion and thus might be promising starting points for

cell engineering to increase specific productivity. Cross-species miRNA microarrays were used as screening tools and quantitative reverse transcription polymerase chain reaction (qRT-PCR) for the confirmation of differentially expressed miRNAs. Samples for comparative physiological analyses are generally taken in the exponential growth phase. However, the cellular transcriptome of mammalian cells is very dynamic during batch cultivation where the conditions change continuously due to nutrient consumption and the accumulation of metabolites (Hernandez Bort et al. 2012; Koh et al. 2009). For this reason, we applied steady-state cultivation using a continuous process (chemostat). This enabled the cultivation of the cells with a defined specific growth rate in a constant environment.

Because miRNAs regulate gene expression via interaction with their target mRNAs, identifying targets is crucial for understanding the biological function of miRNAs. However, although more than 400 expressed mature miRNAs have been identified in CHO cells (Hackl et al. 2012), the exact biological function of most of them in cultivated cells is still largely unknown. To identify potential miRNA-mRNA interactions, miRNA expression data were linked to mRNA expression data from microarray analysis. But due to a lack of reliable computational prediction tools and CHO-specific experimentally validated miRNA target databases, high-throughput miRNA target identification remains a major challenge.

Material and Methods

Cell lines

Recombinant CHO suspension cell lines expressing the 3D6 single chain Fv-Fc fusion antibody (3D6scFv-Fc) and human serum albumin (HSA) with low and high productivity were established as previously described (Maccani et al. 2014). Briefly, protein-free cultivated dihydrofolate reductase deficient CHO cells (DUKX-B11, ATCC CRL-9096) were used as host cell line. After co-transfection with a pCI-neo mammalian expression vector (Promega, Madison, WI, USA) containing the appropriate gene of interest and a second plasmid (p2-dhfr) which contains the dihydrofolate reductase gene, stable recombinant cells were selected in the presence of G418 and the absence of hypoxanthine and thymidine. Productivity was improved by stepwise increase of the methotrexate (MTX) concentration and two steps of subcloning by limiting dilution. Low producing cell lines (CHO 3D6scFv-Fc low producer and CHO HSA low producer) were selected at 0.1 µM MTX and high producing cell lines (CHO 3D6cFv-Fc high producer and CHO HSA high producer) at 0.4

μM MTX. A stable non-producing cell line (CHO empty vector) was established by co-transfection of the host cell line with the empty pCI-neo vector and p2-dhfr. MTX concentration was increased to 0.1 μM.

Steady-state cultivation

Chemostat cultivations were conducted in 800 mL cell culture bioreactors (DS0700TPSS, DASGIP, Jülich, Germany). The inocula were expanded in spinner flasks starting from the working cell bank. Exponentially growing cells from passage six were used for inoculation. The initial cell concentration was 2.5×10^5 cells mL^{-1}. The cultures were maintained at 37°C, pH 7.0, 30% dissolved oxygen and an agitation speed of 80 rpm. The medium was composed of DMEM without glucose and Ham's F12 (1:1) supplemented with 4 mM L-glutamine, 0.25% soy peptone (Quest International, Naarden, The Netherlands), 0.1% Pluronic F68 and a protein free supplement (Polymun Scientific, Klosterneuburg, Austria). After 3 days of batch cultivation, the process was switched to continuous operation for 11 days. Fresh medium was supplied at a constant flow rate to maintain a dilution rate D of 0.5 d^{-1}. The working volume was kept constant at 400 mL using a DASGIP level sensor. Samples for off-line monitoring were taken once a day. D-glucose, L-glutamine, L-glutamate and ammonium concentration were measured with a bioprofile analyzer (BioProfile 100 Plus, Nova Biomedical, Waltham, MA, USA). Cell concentration was determined by counting the nuclei of lysed cells with a Z2 Coulter Counter (Beckman Coulter, Brea, CA, USA). Cell viability was determined by trypan blue exclusion using a hemocytometer.

Productivity determination

The concentrations of the secreted products were determined from the culture supernatants using sandwich ELISA assays as previously described (Maccani et al. 2014). Briefly, 96-well microtiter plates (Nunc MaxiSorp, Thermo Fisher Scientific, Waltham, MA, USA) were coated with 0.33 μg mL^{-1} goat anti-human IgG (γ-chain specific) antibody (I3382, Sigma-Aldrich, St. Louis, MO, USA) to detect 3D6scFv-Fc. Affinity purified 3D6scFv-Fc was used as a standard. Standard and samples were applied in serially two-fold dilutions and captured 3D6scFv-Fc was incubated with 0.5 μg mL^{-1} horseradish peroxidase conjugated goat anti-human IgG (γ-chain specific) antibody (62-8420, Life Technologies, Carlsbad, CA, USA). Staining was conducted using o-phenylenediamine and H_2O_2. The absorption was measured at 492 nm with an infinite M1000 microplate reader (Tecan, Männedorf, Switzerland). HSA

concentrations were determined using the Human Albumin ELISA Quantification Set (E80-129, Bethyl, Montgomery, TX, USA) according to the manufacturer's instructions.

The specific product secretion rate q_P (pg cell^{-1} d^{-1}) during steady-state cultivation was calculated according to equation (1), where D (d^{-1}) represents the dilution rate. VCC (cells mL^{-1}) is the viable cell concentration and C_P (µg mL^{-1}) the product concentration.

$$q_P = D \times \frac{C_P}{VCC} \times 10^6 \quad (1)$$

RNA Isolation

Total RNA samples were isolated from 5×10^6 cells using the Ambion TRI Reagent (Life Technologies, Carlsbad, CA, USA) according to the manufacturer's instructions using chloroform for extraction. Yield and purity were determined using the NanoDrop 1000 sprectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). Only total RNA samples with an A260/A280 ratio between 1.8 and 2.0 and an A260/A230 ratio >2.0 were used in this study. The integrity of the RNA samples was analyzed using the Agilent 2100 Bioanalyser together with the RNA 6000 Nano LabChip kit (Agilent, Santa Clara, CA, USA). The RNA integrity number (RIN) was ≥9.9 for all samples which indicates a very high sample quality.

MicroRNA microarray

Cross-species microRNA microarray assays were conducted as described previously (Hernandez Bort et al. 2012). Briefly, epoxy-coated Nexterion glass slides were spotted with eight replicates of a locked nucleic acid (LNA) probe set consisting of 2367 probes against human, mouse and rat miRNAs based on miRBase 16. Total RNA extracts of three biological replicates per cell line from independent steady-state cultivations were analyzed. Therefore, 800 ng total RNA were hybridized against a common reference (pooled RNA from all samples). To label the miRNAs, the Exiqon Power Labeling Kit (Exiqon, Vedbaek, Denmark) was used according to the manufacturer's instructions. The arrays were hybridized for 16 h at 56°C followed by automated washing and drying with nitrogen using a Tecan HS 400 hybridization station (Tecan, Männedorf, Switzerland). Slides were then scanned at 10 µm resolution and auto-gain settings using a Roche NimbleGen MS200 scanner (Roche NimbleGen, Madison, Wi, USA).

Feature extraction was conducted using the GenePix software (Molecular Devices, Sunnyvale, CA, USA). The LIMMA package of R/Bioconductor was applied for background correction, normalization and statistical analysis as previously described (Hackl et al. 2010). The resulting *p*-values were corrected for multiple testing according to Benjamini and Hochberg (Benjamini and Hochberg 1995). Raw and normalized microarray data have been deposited in NCBI's Gene Expression Omnibus (GEO) database (www.ncbi.nlm.nih.gov/geo/) and are available under accession number GSE57023.
The software Genesis 1.7.6 (Sturn et al. 2002) was used to conduct hierarchical clustering.

mRNA microarray

As microarray platform, the 4×44k design from Agilent (CA, Santa Clara, USA) was chosen. 60-mer oligonucleotide probes were designed based on the published genomic sequence of the CHO-K1 cell line (Xu et al. 2011). The probe-set and array-design (20650 genes, spotted in duplicates) were submitted to the Agilent eArray platform.
Total RNA extracts of three biological replicates per cell line from independent steady-state cultivations were analyzed in duplicates (dye-swap). The Agilent Low Input Quick Amp Labeling Kit was used to generate fluorescent cRNA targets for hybridization with CHO-specific oligonucleotide arrays. Labeling and hybridization were performed according to the manufacturer's instructions. Briefly, 200 ng of total RNA were used for reverse transcription and the subsequent cRNA synthesis and labeling reaction with either cyanine 3-CTP (Cy3) or cyanine 5-CTP (Cy5). After purification of labeled cRNA using the RNeasy Mini Kit (Qiagen, Venlo, The Netherlands), yield and labeling efficiency was determined using the NanoDrop 1000 sprectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). The labeling efficiency was >22 pmol Cy3 or Cy5 per µg cRNA for all samples. The cRNA of the appropriate sample and the common reference (pooled RNA from all samples) were mixed and fragmented using the Agilent Gene Expression Hybridization Kit and transferred to the microarray slide. Hybridization was performed at 65°C for 17 h. After washing, the slides were scanned at 5 µm resolution using an Agilent microarray scanner G2565AB.
The scanned images were processed using the Agilent Feature Extraction 11.0 software. Background correction, normalization and statistical analysis were performed as previously described (Graf et al. 2008). The resulting *p*-values were adjusted for multiple testing using the method of Benjamini and Yekutieli (Reiner et al. 2003).

Quantitative reverse transcription PCR

MicroRNA and mRNA expression were measured using the miScript PCR system (Qiagen, Venlo, The Netherlands) which allows the parallel quantification of mature miRNAs and mRNAs. Total RNA extracts were converted into cDNA using the miScript II RT Kit (Qiagen) according to the manufacturer's instructions. Quantitative real-time PCR (qPCR) was performed on a MiniOpticon real-time PCR detection system (Bio-Rad, Hercules, CA, USA) using the miScript SYBR Green PCR Kit (Qiagen) according to the supplier's manual. To improve the reliability of the assay, the expression of each miRNA was normalized using two internal references (cgr-miR-185-5p and *Actr5*). The miRNA cgr-miR-185-5p was used as an internal control previously (Jadhav et al. 2012) and the mRNA *Actr5* showed very stable expression in the microarray experiment across all CHO cell lines used in this study. Additionally, *Actr5* was described as a suitable internal control gene before (Bahr et al. 2009). For mRNA quantification, *Gapdh* and *Actr5* were used as internal reference genes. A 20 µL qPCR reaction mix contained 10 ng cDNA and the appropriate 10× miScript Primer Assay (Qiagen). All miScript Primer Assays and other primers used in this study are specified in Table S1 (Supplementary material). The PCR was run at 95°C for 15 min and 40 cycles of 94°C for 15 s, 55°C for 30 s and 70°C for 30 s. The specificity of the reactions was verified by analyzing the melting curve immediately after the last amplification cycle. The results were evaluated with the software CFX Manager 3.0 (Bio-Rad). Quantification cycle (C_q) values were determined using the 'Regression'-mode. Three biological replicates (samples from three independent cultivations) were analyzed in technical duplicates (two independent qPCR assays). Relative expression ratios were calculated using the software REST 2009 (Pfaffl et al. 2002). The REST 2009 algorithm uses a statistical randomization test to determine the significance of the expression ratio as well as a complex Taylor series to estimate the standard error (SE).

Target prediction of differentially expressed miRNAs

For miRNA target prediction the miRWalk database (Dweep et al. 2011) was used. Beside the miRWalk algorithm this database includes miRNA-mRNA interactions predicted by DIANA-microT (version 3.0), miRanda (August 2010), miRDB (April 2009), PicTar (March 2007), PITA (August 2008), RNA22 (May 2008), RNAhybrid (version 2.1) and TargetScan (version 5.1). Only interactions predicted by more than half of these programs were considered. The database miRTarBase 4.5 (Hsu et al. 2014) was used to identify potential targets based on

validated miRNA-target interactions in human, mouse or rat. Pearson correlation was analyzed between miRNA expression levels (qRT-PCR data) and mRNA expression levels (microarray data). Student's t-tests was used to identify significantly negatively correlated miRNA-mRNA pairs. The resulting *p*-values were corrected for multiple testing according to Benjamini and Hochberg (Benjamini and Hochberg 1995).

Results

Steady-state cultivation

The CHO cell lines were cultivated in a continuous process (chemostat) to establish steady-state conditions. After a batch phase of three days, the process was switched to continuous operation with a constant dilution rate D of 0.5 d^{-1}. The viable cell concentration remained constant during continuous operation (Fig. 1a–e). Consequently, the specific growth rate μ was equal to the dilution rate D. Also the residual glucose, glutamine, lactate and ammonium reached constant concentrations (Fig. S1a–e, supplementary material), confirming the establishment of steady-state conditions. The specific product secretion rates q_P were considerably (8-fold) different between high and low producers but similar for 3D6scFv-Fc and HSA (Fig. 1f). Samples for miRNA and mRNA expression analysis were harvested on day 14 after more than five volume changes.

The high producing cell lines had been established by increasing the transgene copy number of the low producers in order to reduce effects caused by clonal variation (Maccani et al. 2014). Consequently, observed differences in miRNA and mRNA expression profile between high and low producer should predominantly be relatable to cellular processes that are involved in the biosynthesis of the recombinant protein including transcription, mRNA processing and translation as well as protein processing and secretion. The product mRNA levels perfectly corresponded to the determined increase of q_P between high and low producers (Fig. 2), indicating that there are no limitations at the stage of translation, protein processing or secretion. However, differential gene expression between the high and low producers may reflect the adaptations necessary for the cell to handle the higher recombinant protein load.

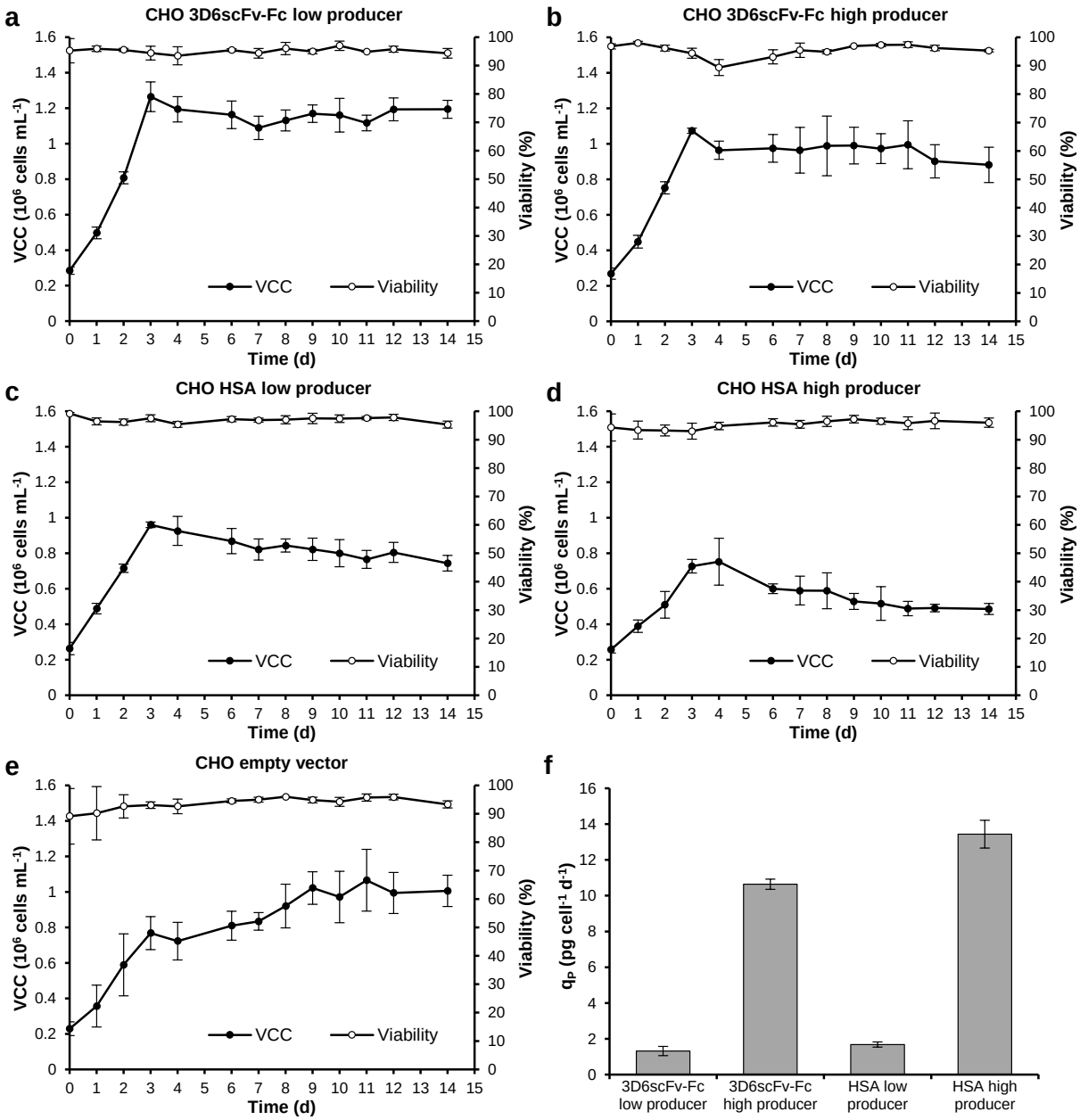


Fig. 1 Time courses of steady-state cultivations. Viable cell concentration and viability of (**a**) CHO 3D6scFv-Fc low producer, (**b**) CHO 3D6scFv-Fc high producer, (**c**) CHO HSA low producer, (**d**) CHO HSA high producer and (**e**) CHO empty vector (non-producer). (**f**) Specific product secretion rate q_P in steady-state. Cells were cultivated in a 0.8 L cell culture bioreactor. After three days of batch cultivation, the process was switched to continuous cultivation (dilution rate D = 0.5 d^{-1}). The culture volume was maintained at a constant level of 400 mL. Data represent mean values of three independent cultivations (error bars: SD)

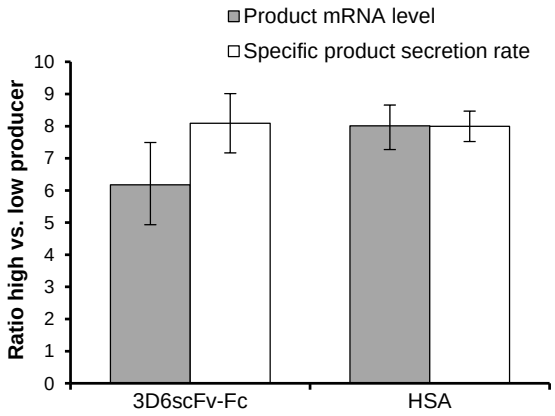


Fig. 2 Correlation of product mRNA level and specific product secretion rate q_P. The diagram shows the ratios of product mRNA and q_P of 3D6scFv-Fc and HSA high producers to low producers. Transcript levels were determined by qRT-PCR. Data represents mean values of three independent steady-state cultivations (error bars: SE)

MicroRNA expression screening by microarray analysis

Expressed and differentially regulated miRNAs were initially identified using a miRNA microarray platform containing probes against 2367 mature miRNAs from human, mouse and rat. In total 320 miRNAs were significantly expressed (signal intensity > background intensity + 2 × standard deviation) in the five CHO cell lines. For these miRNAs, $\log_2$ fold changes against a common reference pool were calculated and the distribution analyzed (Fig. 3a). The results showed increased miRNA levels in the producing cell lines compared to the non-producer, where the shift was statistically significant (Student's t-test, $p < 0.05$) for 3D6scFv-Fc low producer and 3D6scFv-Fc high producer. Comparing high producers with low producers as well as producers with non-producer, a total of 83 non-redundant miRNAs were significantly differentially expressed (adj. $p < 0.05$ and fold change > 1.5) in at least one of the comparisons. These miRNAs were analyzed using hierarchical clustering (Fig. 3b). The heat map shows that 3D6scFv-Fv expression predominantly led to an up-regulation of miRNA expression (Cluster D and E). In contrast, a large fraction of miRNAs was down-regulated in both HSA producers (Cluster A, B and C). Especially the miRNAs of cluster A were highly affected by HSA expression. However, comparing the high producers with the low producers, no significantly differentially expressed miRNA (adj. $p < 0.05$ and fold change > 1.5) was commonly up- or down-regulated for both model proteins (Fig. 3c). Interestingly, more miRNAs were significantly down-regulated than up-regulated between the individual high and low producers, although we observed a global increase of miRNA levels in the

producing cell lines relative to the non-producer. Between high producers and non-producer five significantly differentially expressed miRNAs were commonly up-regulated and one was commonly down-regulated (Fig. 3d).

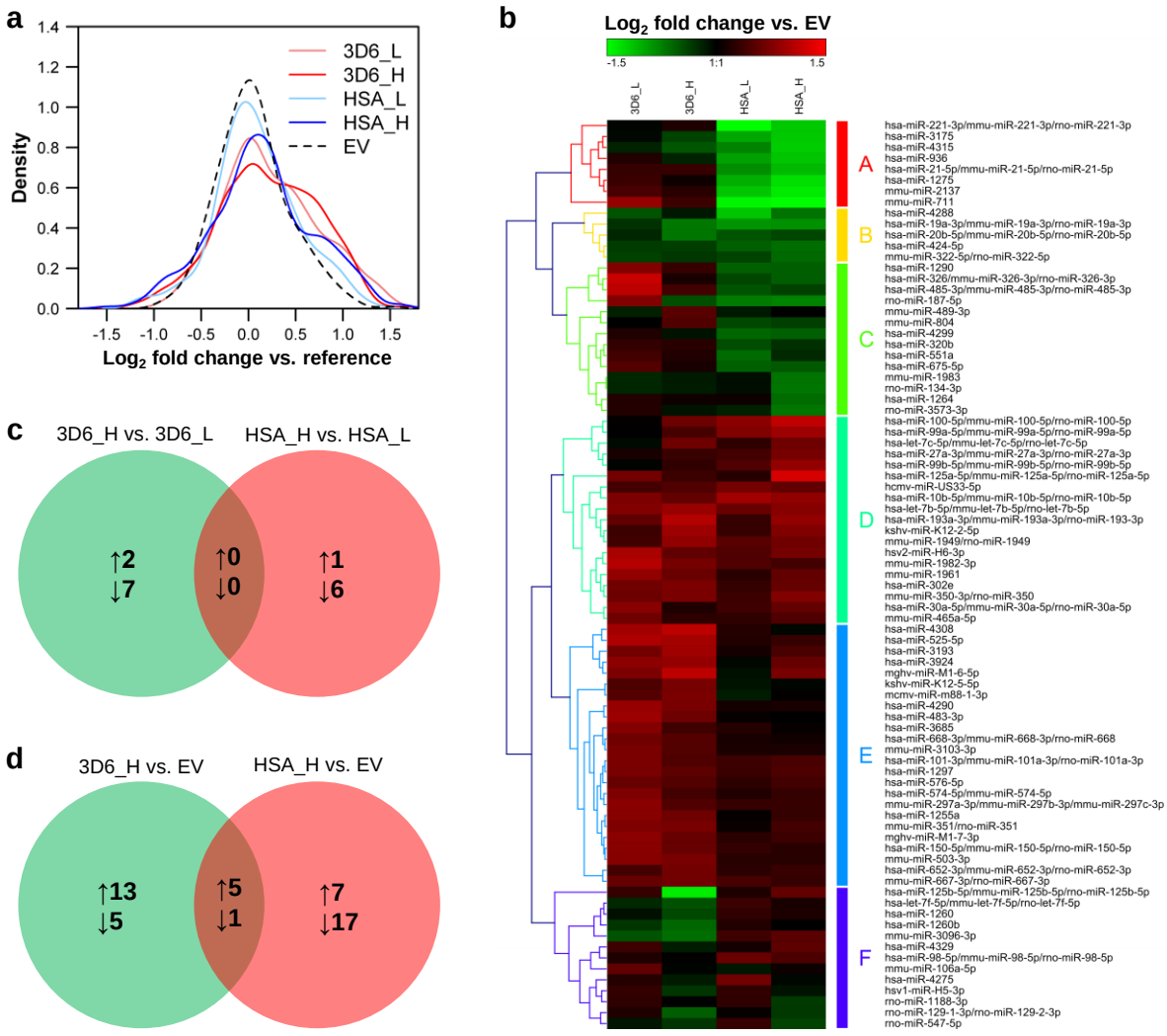


Fig. 3 Comparative microRNA profiling using microarray analysis. Total RNA samples of five CHO cells lines from steady-state cultivations (n = 3) were analyzed. 3D6_L, CHO 3D6scFv-Fc low producer; 3D6_H, CHO 3D6scFv-Fc high producer; HSA_L, CHO HSA low producer; HSA_H, CHO HSA high producer; EV, CHO empty vector (non-producer). (**a**) Density plot of the $\log_2$ fold change miRNA expression between each cell line and a common reference pool (mean values, n = 3). (**b**) Hierarchical clustering of 83 significantly differentially expressed mature miRNAs (adj. $p < 0.05$ and fold change > 1.5) based on $\log_2$ fold changes between producers and non-producer. Commonly and exclusively up- or down-regulated miRNAs were determined using Venn diagrams. The number of significantly differentially expressed miRNAs of (**c**) 3D6scFv-Fc high producer versus 3D6scFv-Fc low producer and HSA high producer versus HSA low producer and (**d**) 3D6scFv-Fc high producer versus non-producer and HSA high producer versus non-producer are illustrated. ↑, up-regulated; ↓, down-regulated

Differential miRNA expression by qRT-PCR

A total of 14 mature miRNAs from the microarray study were selected for qRT-PCR. The selection was based on the observed most significantly differentially expressed miRNAs between high and low producers as well as between the individual producers and the non-producer including the miRNAs found to be commonly up- or down-regulated between both high producers and the non-producer. Two of them were human-specific miRNAs (hsa-miR-936 and hsa-miR-3175), one was mouse-specific (mmu-miR-711) and the others showed a 100% sequence identity between the known mature mouse, human, rat and Chinese hamster miRNA according to miRBase 20 (Kozomara and Griffiths-Jones 2014). However, only the 11 conserved miRNAs could be reliably detected by qRT-PCR (Fig. 4). The mature sequences of hsa-miR-936, hsa-miR-3175 and mmu-miR-711 could also not be found in the genome of CHO-K1 (Xu et al. 2011) and the Chinese hamster (Brinkrolf et al. 2013; Lewis et al. 2013) by BLAST search (100% identity) using the Chinese hamster genome database (Hammond et al. 2012), indicating that those Chinese hamster RNAs leading to a signal on the microarrays are different from the human and mouse miRNAs. In addition, there are no similar Chinese hamster miRNAs which are included in miRBase 20 that could have cross-hybridized to the microarray probes of hsa-miR-936, hsa-miR-3175 and mmu-miR-711.

The miRNAs let-7b-5p and let-7c-5p were up-regulated in all producing cell lines compared to the non-producer and the expression levels correlated with the productivity. Comparing the 3D6scFv-Fc high and low producer, miR-99a-5p was significantly ($p < 0.005$ and fold change > 1.5) up-regulated and miR-100-5p, miR-125b-5p and miR-19a-3p significantly down-regulated. Of these, miR-100-5p and miR-125b-5p showed the inverse effect between HSA high and low producer, where they were significantly up-regulated.

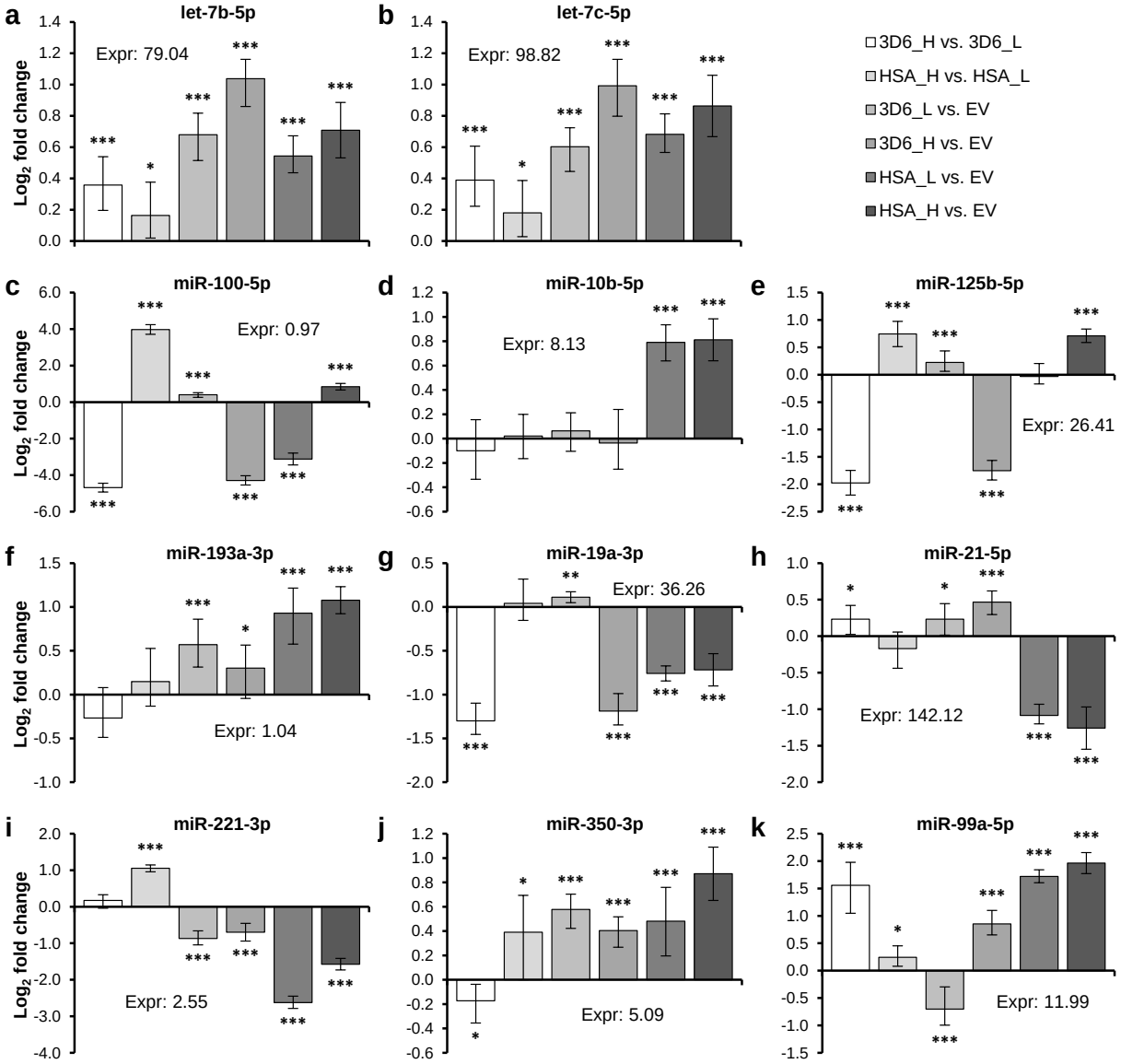


Fig. 4 Differentially expressed miRNAs determined by qRT-PCR. 3D6_H, CHO 3D6scFv-Fc high producer; 3D6_L, CHO 3D6scFv-Fc low producer; HSA_H, CHO HSA high producer; HSA_L, CHO HSA low producer; EV, CHO empty vector (non-producer). qRT-PCR data were normalized using two endogenous controls (miR-185-5p and *Actr5*). The software REST 2009 was used to calculate relative expression ratios and for statistical analysis (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$). Data represent mean values of three independent steady-state cultivations (error bars: SE)

Identification of potential miRNA-mRNA interactions

In order to generate hypotheses about the biological function of the 11 conserved miRNAs analyzed by qRT-PCR, potential target mRNAs were determined based on computational target prediction algorithms and experimentally validated targets, combined with the identification of miRNA-mRNA pairs that show a negatively correlated expression profile.

Computational miRNA target prediction was conducted using miRWalk (Dweep et al. 2011), DIANA-microT (Maragkakis et al. 2009), miRanda (John et al. 2004), miRDB (Wang 2008), PicTar (Krek et al. 2005), PITA (Kertesz et al. 2007), RNA22 (Miranda et al. 2006), RNAhybrid (Rehmsmeier et al. 2004) and TargetScan (Lewis et al. 2005). To reduce the false positive rate, only mRNAs predicted by five or more of these nine algorithms were considered. In a second approach, potential targets were obtained from miRTarBase 4.5 using the experimentally validated miRNA-mRNA interactions in human, mouse or rat.

The mRNA microarray experiment revealed a total of 2842 genes which were significantly differentially expressed (adj. $p < 0.05$ and fold change > 1.5) in at least one comparison between the five cell lines analyzed in this study. Pearson correlation coefficients (PCC) were computed for the 11 differentially expressed miRNAs (expression levels from qRT-PCR) and the 2842 differentially expressed mRNAs (expression levels from microarray experiment). Identified potential miRNA-mRNA interactions that showed a negative correlation (PCC < -0.5 and adj. $p < 0.05$) are listed in Table 1. Fisher's exact test was used to analyze whether negatively correlated targets are significantly ($p < 0.05$) enriched. For the computationally predicted targets, no significant enrichment was observed (Table 2). Regarding the experimentally validated targets, negatively correlated ones were significantly enriched for miR-21-5p and miR-99a-5p (Table 2). Kernel density plots were computed to visualize and compare the distribution of the PCCs for validated targets, predicted targets and total differentially expressed mRNAs (Fig. S2, supplementary material). Generally, a shift to negative PCCs was observed. However, the shifts were only statistical significant (Student's t-test, $p < 0.05$) for the validated targets of miR-10b-5p, miR-21-5p and miR-125b-5p as well as the predicted targets of let-7c-5p. Hence, these results indicate a considerable degree of false positives within the discovered miRNA-mRNA interactions.

Table 1 Identified negatively correlated potential targets of differentially expressed miRNAs

Mature miRNA	Potential target mRNAs[a]
let-7b-5p	*Adam8*, *Ammecr1l*, *Atad3a*, *Cpsf3l*, *Ddx26b*, *Ddx41*, ***Dmd***, *Dnmbp*, *Eef1e1*, *Fam49b*, ***Gltpd1***, *Iars*, ***Igf2bp3***, *Lingo1*, ***Lman2***, ***Mars2***, *Nid1*, *Prss22*, *Ptgs2*, *Slc30a4*, ***Syncrip***, *Taf5*, ***Tmem65***, *Tmprss11f*, *Txndc5*, *Zadh2*, *Zbtb5*
let-7c-5p	*Adam8*, *Ammecr1l*, *Atad3a*, *Bzw1*, *Col4a1*, *Ddx18*, *Ddx26b*, *Fkbp10*, *Gltpd1*, *Golt1b*, *Iars*, *Igf2bp3*, *Jarid2*, *Lingo1*, *Lman2*, *Mars2*, *Pld3*, *Plxna2*, *Prss22*, *Rb1*, *Slc30a4*, *Slk*, *Stk24*, *Syncrip*, *Tmprss11f*, *Txndc5*, *Vps25*, *Zbtb5*
miR-100-5p	-
miR-10b-5p	*Acly*, *Ap3m1*, *Arhgap18*, *Col4a1*, *Ctdspl*, *Eif1*, *Glod4*, *Gpc1*, *Hivep2*, *Idh3a*, *Igf2r*, *Igfbp4*, *Klhdc7a*, *Lrrc16a*, *Lrrc59*, *Mboat1*, *Nkiras2*, *Ormdl1*, *Ppp1r9b*, *Rap2a*, *Rhobtb2*, *Sdc1*, *Slc25a30*, *Stk4*, *Tox4*, *Txndc16*, ***Ube2z***
miR-125b-5p	*6430548M08Rik*, *Ak3*, *Akap1*, *Ankrd13b*, *Arrb1*, *BC003266*, ***Cd320***, *Cln6*, *Cspg4*, *Cyp2c55*, *Cyyr1*, *Dusp3*, *Dynlt3*, *Ebpl*, *Fbn1*, *Fbxw4*, *Fgfr2*, *Gbf1*, *Ghdc*, *Gpc6*, *Gsn*, *Hspd1*, *Icmt*, *Ier3*, *Ilvbl*, ***Jub***, *Ldb1*, ***Lss***, *M6pr*, *Mamdc2*, *Map3k1*, *Mgat5*, *Myt1*, *Ngly1*, *Nup50*, *Osgepl1*, *Pde1a*, *Phex*, *Ppwd1*, *Rere*, *Sept3*, *Slc27a6*, *Snx8*, *St8sia4*, *Stat5b*, *Taf15*, *Thop1*, *Tmem180*, *Tmem201*, *Tnfsf4*, *Tspan9*, *Zmym3*
miR-193a-3p	*Abcc3*, *Acpl2*, *Ap2b1*, *Avpi1*, *Bicd2*, *Ccdc134*, *Cyb561d2*, *Dnajc7*, *Faf2*, *Fzd4*, *Gnat2*, *Igf2bp3*, *Igf2r*, *Ing5*, *Kdelc1*, *Kras*, *Lamc2*, *Lrrc16a*, *Mmp14*, *Nkiras2*, *Npepps*, *Phf21b*, *Slc4a3*, *Slmap*, *Spsb4*, *Tmed3*, *Vps37b*
miR-19a-3p	*Abhd10*, *Ahrr*, *Atp10a*, *Bmp2k*, *Depdc1b*, *Dsel*, *Icmt*, *Mdfic*, *Mid1ip1*, *Pde5a*, *Pdik1l*, *Pls3*, *Rras2*, *Snx7*, *Stx12*, *Timp2*, *Tmem50a*, *Zeb2*
miR-21-5p	*Aftph*, *Alx1*, *Ank2*, *Ankrd28*, *Arhgap24*, ***Atp11b***, *B3galnt1*, *Bmpr1b*, *Boc*, *Ccdc117*, *Cd44*, *Cryab*, *Ctdsp2*, *Dmd*, *Dock4*, *Dse*, *Dync1li2*, *Elf2*, ***Elovl7***, *Entpd5*, *Fbxl17*, *Fubp1*, *Glcci1*, *Gng12*, *Gpam*, *Grsf1*, *Hoxa9*, *Icam1*, ***Mbnl1***, *Mthfd2*, *Nbea*, *Nek1*, ***Nfib***, *Nkiras1*, ***Pdcd4***, *Phf20l1*, *Pias3*, *Pkd2*, *Postn*, *Ppap2a*, *Ptx3*, *Pura*, *Rabgap1l*, *Rbms3*, *Rnf11*, *Rnf167*, *Rpa2*, *Rufy3*, *Sash1*, *Smap2*, ***Srpk2***, ***Taf5***, *Timp2*, ***Trim33***, *Ttc33*, *Uso1*, *Wwc2*, *Zbtb38*, *Zfp110*, *Zfp112*, *Zfp367*
miR-221-3p	*2900011O08Rik*, *Ank2*, *Ankrd28*, *Bmp2k*, *Bmpr1a*, *Capn7*, *Carhsp1*, *Casp9*, *Ccnd2*, *Cd44*, ***Cdkn1b***, *Cxcr7*, *Eaf1*, *Eif4g3*, *Elavl2*, *Eya4*, *Fhl1*, *Figf*, *Glud1*, *Gnptab*, *Gpbp1*, *Icam1*, *Mbnl1*, *Mdfic*, *Nkiras1*, *Nt5dc2*, *Pak1*, ***Pdik1l***, *Phf21a*, *Pkd2*, *Plscr4*, *Ptx3*, *Rfx7*, *Rpl15*, *Sema3b*, *Sema3e*, *Sh3d19*, *Shmt2*, *Slc25a12*, *Slc33a1*, *Slc4a7*, *Sqstm1*, *Ssbp2*, *Stmn1*, *Tapbp*, *Tmem140*, *Tmem176b*, *Wwc2*
miR-350-3p	-
miR-99a-5p	*Ap2b1*, *Apex1*, *Arhgap22*, *Col4a1*, *Ctdspl*, *Ddhd1*, *Ddx18*, *Lman2*, *Nfe2l1*, *Ormdl1*, *Rb1*, *Scpep1*, *Serpine1*, *Smarca5*, *Sucla2*, *Trib1*, *Zfp689*

[a] Computationally predicted miRNA targets, experimentally validated miRNA targets in human, mouse or rat (underlined) or determined by both methods (bold)

Table 2 Enrichment analysis of negatively correlated miRNA targets

miRNA	Number of negatively correlated differentially expressed genes[a]	Number of differentially expressed targets		Number of negatively correlated differentially expressed targets		Odds ratio (OR)[b]		*p*-value (Fisher's exact test)[c]	
		Predicted	Validated	Predicted	Validated	Predicted targets	Validated targets	Predicted targets	Validated targets
let-7b-5p	237	169	205	19	16	1.392	0.931	0.200	0.896
let-7c-5p	314	162	19	24	5	1.400	2.874	0.158	0.052
miR-100-5p	11	17	44	0	0	-	-	-	-
miR-10b-5p	860	57	28	19	12	1.152	1.728	0.663	0.153
miR-125b-5p	514	316	56	50	11	0.851	1.107	0.353	0.727
miR-193a-3p	710	93	0	28	0	1.293	-	0.275	-
miR-19a-3p	243	200	10	18	0	1.058	-	0.794	-
miR-21-5p	825	122	115	41	46	1.237	1.630	0.309	0.016
miR-221-3p	647	159	42	47	14	1.423	1.696	0.053	0.137
miR-350-3p	0	264	0	0	0	-	-	-	-
miR-99a-5p	925	10	24	6	13	3.107	2.448	0.088	0.030

[a] 2842 differentially expressed genes in total

[b] Indicates the degree of enrichment/depletion. OR > 1, negatively correlated differentially expressed targets are overrepresented. OR < 1, negatively correlated differentially expressed targets are underrepresented

[c] Significance of enrichment/depletion

Discussion

Comparable conditions by steady-state cultivation

A prerequisite for a comparative physiological analysis of different cell lines is the generation of samples under comparable and defined conditions in order to obtain reproducible and meaningful information. In simple batch cultures the physico-chemical conditions are very dynamic and have a considerable impact on the cell's transcriptome (Hernandez Bort et al. 2012; Koh et al. 2009). Koh et al. observed significant changes in miRNA expression even within the exponential growth phase of batch cultivated HEK-293 cells. This clearly shows that in batch cultivations the time of sampling is crucial for the outcome of omics studies that compare different cell lines. Consequently the chemostat with its defined and constant environment is the ideal set-up for such experiments (Hoskisson and Hobbs 2005). Chemostat cultures have been used to grow CHO cells under steady-state condition for more than two decades (Hayter et al. 1992). However, as the standard industrial bioprocess is fed-batch, steady-state cultivation was not applied in recent CHO-omics studies.

Product dependency of miRNA expression profile

The most striking observation in our study was the divergence in differentially regulated miRNAs between the cell lines producing different proteins. We observed a general increase of miRNA expression levels in the producing cell lines, which indicates that miRNAs play an important role in the regulation of processes involved or caused by recombinant protein synthesis and secretion. Several miRNAs were differentially expressed comparing high, low and non-producing CHO cell lines. However, significant differences in miRNA expression were predominantly seen between producers and non-producers as well as between 3D6scFv-Fc and HSA producers, rather than between high and low producers. Furthermore, no significantly differentially expressed miRNA was commonly up- or down-regulated comparing high and low producers for both model proteins. These results suggest that the reaction of CHO cells to recombinant protein expression strongly depends on the particular product, which would also explain the low level of consensus observed between previously published studies investigating the transcriptome of different CHO production cell lines in relation to high productivity (Vishwanathan et al. 2014). Different proteins require specific cellular aid with folding, glycosylation and bonding depending on their structure. The exerted stress may well initiate differential responses and consequently affect the regulation of both miRNAs and mRNAs that provide the cells with the capacity to handle the various types of stress. As the mRNA expression of the product genes correlated well to the cell specific productivity, it appears that the cellular protein production machinery was able to adapt to these individual requirements without running into limitation. Nevertheless, the contribution of effects caused by clonal variation is unknown and cannot be neglected.

Impact of confirmed miRNAs

For those differentially regulated miRNAs that were confirmed by qRT-PCR, some consensus could be found in similar previously published studies. In an earlier work, we compared the miRNA expression of an Epo-Fc and an antibody producing cell line to the respective parental CHO cell line (Hackl et al. 2011). We observed an up-regulation of miR-10b-5p and down-regulation of miR-21-5p in recombinant cell lines, which is in compliance with our results of the HSA producers, but not for the 3D6scFv-Fv producers. Additionally, we could already show that miR-21-5p overexpression reduces specific productivity (Jadhav et al. 2012). Consequently, this suggests that a knockdown of miR-21-5p could increase specific productivity. As one of the best studied miRNAs, miR-21-5p has been linked to cell

proliferation, apoptosis and migration (Krichevsky and Gabriely 2009) and even cellular longevity (Dellago et al. 2013). Like miR-21-5p, miR-10b-5p was described as an oncogenic miRNA with a pro-proliferative and anti-apoptotic function (Lin et al. 2012).

In another study, Lin et al. (2011) profiled miRNA expression in four recombinant CHO cell lines expressing the same human IgG and compared them with the parental DG44 cell line. In compliance with our results, they found miR-221-3p being significantly down-regulated in the recombinant cell lines. They also observed a down-regulation of miR-125b-5p in two clones, as observed in the 3D6scFv-Fc high producer in our study. However, they detected opposed effects for miR-19a-3p and let-7b-5p. In human hepatocellular carcinoma cells, miR-221-3p was found to control cyclin-dependent kinase inhibitor 1B ($p27^{Kip1}$) and cyclin-dependent kinase inhibitor 1C ($p57^{Kip2}$) expression (Fornari et al. 2008). Induction of G1-specific growth arrest by conditional overexpression of $p27^{Kip1}$ resulted in increased specific productivity in CHO (Meents et al. 2002), however, transient overexpression of miR-221 had no significant effect on growth or productivity (Jadhav et al. 2012). This might indicate that miR-221 expression changes in response to cellular requirements, but forced manipulation of its expression does not by itself change the phenotype. miR-125b-5p can act as tumor suppressor or as oncogene and it was shown to promote apoptosis by suppressing the expression of Bcl-2 family members (Gong et al. 2013).

All miRNAs described above, have been associated with cell growth and/or apoptosis, so that their relation to productivity may only be indirect and therefore difficult to interpret. Hence, further evaluation of the identified target miRNAs is required.

The challenge of high-throughput miRNA target identification

Identifying the biological function of a miRNA is still a major challenge. Due to the complexity and diversity of miRNA-mRNA target interactions, functional screenings using biological methods are very labor-intensive and time-consuming. Hence, reliable computational tools for the prediction of interactions between miRNAs and target mRNAs would be a huge benefit. But the prediction of targets is very challenging as miRNAs recognize specific sequences with only partial complementarity. However, although perfect pairing between nucleotides 2–7 (seed region) and the target site is the most common motif in animals (Pasquinelli 2012), also imperfect seed pairing compensated by extensive pairing of the 3’ end (Vella et al. 2004) and centered pairing (Shin et al. 2010) have been described. Numerous algorithms based on seed pairing, evolutionary conserved sites, secondary structure of the 3’ UTR and thermodynamic calculations have been developed for computational target

prediction. Many of these algorithms consider cross-species conservation to reduce false positive rates (Maziere and Enright 2007). However, by applying a novel biochemical approach known as high-throughput sequencing of RNA isolated by crosslinking immunoprecipitation (HITS-CLIP), it was shown that a considerable fraction (40%) of all functional target sites is not conserved (Ellwanger et al. 2011). The lack of agreement between the results of different computational methods as well as the high false-positive and false-negative rates clearly show the complexity of target prediction in mammalian cells (Liu et al. 2014; Ritchie et al. 2009). In CHO research the reliability of the currently available target prediction tools is additionally impaired as the CHO or Chinese Hamster genome has not been included in any of them yet. Hence, the results rely on an assumed high degree of conservation of miRNA and target mRNA interactions between Chinese hamster and mouse, rat or human.

In mammalian cells, various studies suggest that miRNAs predominantly act by decreasing target mRNA levels (Baek et al. 2008; Guo et al. 2010; Hendrickson et al. 2009). This supports the frequently used approach of correlating miRNA and mRNA expression levels to identify negatively correlated pairs, which we also applied here. We analyzed whether negatively correlated miRNA-mRNA pairs are enriched within computationally predicted and experimentally validated targets. However, a statistically significant enrichment could only be observed within the validated targets for two of the eleven analyzed miRNAs. This approach also assumes that the target mRNA is predominantly regulated by the miRNA and not at a different stage of gene expression. In addition, it has been shown that most miRNAs are only discernibly active above a certain expression level (Mullokandov et al. 2012). Furthermore, there are several factors that can affect miRNA binding efficiency to a specific target mRNA, including competition with RNA-binding proteins (Kedde et al. 2007), miRISC cofactors (Neumüller et al. 2008) or modification of Argonaute proteins (Johnston and Hutvagner 2011). In addition, mRNAs can contain multiple target sites for a single miRNA and also target sites for several miRNAs which suggests even more complex regulatory mechanisms by miRNAs (Liu et al. 2014). Together this indicates that many targets most likely remain undiscovered by the simple assumption of an inverse relationship between the expression levels of a miRNA and its target mRNAs.

In conclusion, this is the first report of miRNA expression data of recombinant CHO cell lines cultivated under steady-state conditions. Cell lines that express heterologous proteins appear to have higher levels of mature miRNAs in general, which suggests that miRNAs play a

crucial role in recombinant CHO cell lines. However, comparing the miRNA expression profiles of different CHO cell lines, both from this study and published results, revealed little consensus. This indicates that the reaction of CHO cells to the overexpression of heterologous proteins is strongly protein and/or clone dependent. Hence, cell engineering approaches to improve recombinant protein production may also be product and/or clone-specific and not generally applicable.

Identifying miRNAs and their target mRNAs is crucial for a better understanding of biological processes in CHO cells and there is no doubt that high-throughput miRNA and mRNA profiling can deliver valuable information. Additional results from proteome analyses would also be beneficial to obtain a comprehensive picture (Baek et al. 2008). As the regulatory mechanisms of miRNAs are very complex and today's computational target prediction tools are inefficient, it is still a major challenge to retrieve correct and meaningful results from high-throughput omics data. Hence, reliable computational miRNA target prediction tools that also include the CHO genome are urgently needed. This also requires thoroughly defined 3' UTR boundaries in the CHO genome as this highly influences the outcome of miRNA target prediction (Ritchie et al. 2009). In addition, experimentally validated miRNA targets in CHO cells need to be collected in order to finally facilitate functional analysis of high-throughput miRNA expression data.

Acknowledgments

This work has been supported by the Austrian Center of Industrial Biotechnology (ACIB) through the COMET-Funding Program managed by the Austrian Research Promotion Agency FFG with support of the Federal Ministry of Science, Research and Economy (BMWFW), the Federal Ministry of Traffic, Innovation and Technology (bmvit), the Styrian Business Promotion Agency SFG, the Standortagentur Tirol and ZIT - Technology Agency of the City of Vienna. Andreas Maccani was also supported by the PhD program "BioToP – Biomolecular Technology of Proteins" (Austrian Science Fund, FWF Project W1224).

We thank Stefan Bauernfried for his excellent assistance with the fermentations of the CHO cell lines.

Conflict of interest statement

The authors declare no commercial or financial conflict of interest.

References

Baek D, Villen J, Shin C, Camargo FD, Gygi SP, Bartel DP (2008) The impact of microRNAs on protein output. Nature 455(7209):64-71 doi:10.1038/nature07242

Bahr SM, Borgschulte T, Kayser KJ, Lin N (2009) Using microarray technology to select housekeeping genes in Chinese hamster ovary cells. Biotechnol Bioeng 104(5):1041-1046 doi:10.1002/bit.22452

Barron N, Kumar N, Sanchez N, Doolan P, Clarke C, Meleady P, O'Sullivan F, Clynes M (2011) Engineering CHO cell growth and recombinant protein productivity by overexpression of miR-7. J Biotechnol 151(2):204-211 doi:10.1016/j.jbiotec.2010.12.005

Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116(2):281-297 doi:10.1016/S0092-8674(04)00045-5

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate - a practical and powerful approach to multiple testing. J Roy Stat Soc B Met 57(1):289-300 doi:10.2307/2346101

Brinkrolf K, Rupp O, Laux H, Kollin F, Ernst W, Linke B, Kofler R, Romand S, Hesse F, Budach WE, Galosy S, Müller D, Noll T, Wienberg J, Jostock T, Leonard M, Grillari J, Tauch A, Goesmann A, Helk B, Mott JE, Pühler A, Borth N (2013) Chinese hamster genome sequenced from sorted chromosomes. Nat Biotechnol 31(8):694-695 doi:10.1038/nbt.2645

Dellago H, Preschitz-Kammerhofer B, Terlecki-Zaniewicz L, Schreiner C, Fortschegger K, Chang MW, Hackl M, Monteforte R, Kühnel H, Schosserer M, Gruber F, Tschachler E, Scheideler M, Grillari-Voglauer R, Grillari J, Wieser M (2013) High levels of oncomiR-21 contribute to the senescence-induced growth arrest in normal human cells and its knock-down increases the replicative lifespan. Aging Cell 12(3):446-458 doi:10.1111/acel.12069

Druz A, Son YJ, Betenbaugh M, Shiloach J (2013) Stable inhibition of mmu-miR-466h-5p improves apoptosis resistance and protein production in CHO cells. Metab Eng 16:87-94 doi:10.1016/j.ymben.2012.12.004

Dweep H, Sticht C, Pandey P, Gretz N (2011) miRWalk--database: prediction of possible miRNA binding sites by "walking" the genes of three genomes. J Biomed Inform 44(5):839-847 doi:10.1016/j.jbi.2011.05.002

Ellwanger DC, Büttner FA, Mewes HW, Stümpflen V (2011) The sufficient minimal set of miRNA seed types. Bioinformatics 27(10):1346-1350 doi:10.1093/bioinformatics/btr149

Fornari F, Gramantieri L, Ferracin M, Veronese A, Sabbioni S, Calin GA, Grazi GL, Giovannini C, Croce CM, Bolondi L, Negrini M (2008) MiR-221 controls CDKN1C/p57 and CDKN1B/p27 expression in human hepatocellular carcinoma. Oncogene 27(43):5651-5661 doi:10.1038/onc.2008.178

Gong J, Zhang JP, Li B, Zeng C, You K, Chen MX, Yuan Y, Zhuang SM (2013) MicroRNA-125b promotes apoptosis by regulating the expression of Mcl-1, Bcl-w and IL-6R. Oncogene 32(25):3071-3079 doi:10.1038/onc.2012.318

Graf A, Gasser B, Dragosits M, Sauer M, Leparc GG, Tüchler T, Kreil DP, Mattanovich D (2008) Novel insights into the unfolded protein response using *Pichia pastoris* specific DNA microarrays. BMC Genomics 9:390 doi:10.1186/1471-2164-9-390

Guo H, Ingolia NT, Weissman JS, Bartel DP (2010) Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature 466(7308):835-840 doi:10.1038/nature09267

Hacker DL, De Jesus M, Wurm FM (2009) 25 years of recombinant proteins from reactor-grown cells - where do we go from here? Biotechnol Adv 27(6):1023-1027 doi:10.1016/j.biotechadv.2009.05.008

Hackl M, Brunner S, Fortschegger K, Schreiner C, Micutkova L, Mück C, Laschober GT, Lepperdinger G, Sampson N, Berger P, Herndler-Brandstetter D, Wieser M, Kühnel H, Strasser A, Rinnerthaler M, Breitenbach M, Mildner M, Eckhart L, Tschachler E, Trost A, Bauer JW, Papak C, Trajanoski Z, Scheideler M, Grillari-Voglauer R, Grubeck-Loebenstein B, Jansen-Dürr P, Grillari J (2010) miR-17, miR-19b, miR-20a, and miR-106a are down-regulated in human aging. Aging Cell 9(2):291-296 doi:10.1111/j.1474-9726.2010.00549.x

Hackl M, Jakobi T, Blom J, Doppmeier D, Brinkrolf K, Szczepanowski R, Bernhart SH, Höner zu Siederdissen C, Hernandez Bort JA, Wieser M, Kunert R, Jeffs S, Hofacker IL, Goesmann A, Pühler A, Borth N, Grillari J (2011) Next-generation sequencing of the Chinese hamster ovary microRNA transcriptome: Identification, annotation and profiling of microRNAs as targets for cellular engineering. J Biotechnol 153(1-2):62-75 doi:10.1016/j.jbiotec.2011.02.011

Hackl M, Jadhav V, Jakobi T, Rupp O, Brinkrolf K, Goesmann A, Pühler A, Noll T, Borth N, Grillari J (2012) Computational identification of microRNA gene loci and precursor microRNA sequences in CHO cell lines. J Biotechnol 158(3):151-155 doi:10.1016/j.jbiotec.2012.01.019

Hammond S, Kaplarevic M, Borth N, Betenbaugh MJ, Lee KH (2012) Chinese hamster genome database: an online resource for the CHO community at www.CHOgenome.org. Biotechnol Bioeng 109(6):1353-1356 doi:10.1002/bit.24374

Hayter PM, Curling EM, Baines AJ, Jenkins N, Salmon I, Strange PG, Tong JM, Bull AT (1992) Glucose-limited chemostat culture of Chinese hamster ovary cells producing recombinant human interferon-gamma. Biotechnol Bioeng 39(3):327-335 doi:10.1002/bit.260390311

He L, Hannon GJ (2004) MicroRNAs: small RNAs with a big role in gene regulation. Nat Rev Genet 5(7):522-531 doi:10.1038/nrg1379

Hendrickson DG, Hogan DJ, McCullough HL, Myers JW, Herschlag D, Ferrell JE, Brown PO (2009) Concordant regulation of translation and mRNA abundance for hundreds of targets of a human microRNA. PLoS Biol 7(11):e1000238 doi:10.1371/journal.pbio.1000238

Hernandez Bort JA, Hackl M, Höflmayer H, Jadhav V, Harreither E, Kumar N, Ernst W, Grillari J, Borth N (2012) Dynamic mRNA and miRNA profiling of CHO-K1 suspension cell cultures. Biotechnol J 7(4):500-515 doi:10.1002/biot.201100143

Hobert O (2008) Gene regulation by transcription factors and microRNAs. Science 319(5871):1785-1786 doi:10.1126/science.1151651

Hoskisson PA, Hobbs G (2005) Continuous culture--making a comeback? Microbiology 151(Pt 10):3153-3159 doi:10.1099/mic.0.27924-0

Hsu SD, Tseng YT, Shrestha S, Lin YL, Khaleel A, Chou CH, Chu CF, Huang HY, Lin CM, Ho SY, Jian TY, Lin FM, Chang TH, Weng SL, Liao KW, Liao IE, Liu CC, Huang HD (2014) miRTarBase update 2014: an information resource for experimentally validated miRNA-target interactions. Nucleic Acids Res 42(Database issue):D78-D85 doi:10.1093/nar/gkt1266

Jadhav V, Hackl M, Bort JA, Wieser M, Harreither E, Kunert R, Borth N, Grillari J (2012) A screening method to assess biological effects of microRNA overexpression in Chinese hamster ovary cells. Biotechnol Bioeng 109(6):1376-1385 doi:10.1002/bit.24490

Jadhav V, Hackl M, Klanert G, Bort JA, Kunert R, Borth N, Grillari J (2014) Stable overexpression of miR-17 enhances recombinant protein production of CHO cells. J Biotechnol 175:38-44 doi:10.1016/j.jbiotec.2014.01.032

John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS (2004) Human MicroRNA targets. PLoS Biol 2(11):e363 doi:10.1371/journal.pbio.0020363

Johnston M, Hutvagner G (2011) Posttranslational modification of Argonautes and their role in small RNA-mediated gene regulation. Silence 2:5 doi:10.1186/1758-907X-2-5

Kedde M, Strasser MJ, Boldajipour B, Oude Vrielink JA, Slanchev K, le Sage C, Nagel R, Voorhoeve PM, van Duijse J, Orom UA, Lund AH, Perrakis A, Raz E, Agami R (2007) RNA-binding protein Dnd1 inhibits microRNA access to target mRNA. Cell 131(7):1273-1286 doi:10.1016/j.cell.2007.11.034

Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E (2007) The role of site accessibility in microRNA target recognition. Nat Genet 39(10):1278-1284 doi:10.1038/ng2135

Kim JY, Kim YG, Lee GM (2012) CHO cells in biotechnology for production of recombinant proteins: current state and further potential. Appl Microbiol Biotechnol 93(3):917-930 doi:10.1007/s00253-011-3758-5

Koh TC, Lee YY, Chang SQ, Nissom PM (2009) Identification and expression analysis of miRNAs during batch culture of HEK-293 cells. J Biotechnol 140(3-4):149-155 doi:10.1016/j.jbiotec.2009.01.021

Kozomara A, Griffiths-Jones S (2014) miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res 42(Database issue):D68-D73 doi:10.1093/nar/gkt1181

Krek A, Grün D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, da Piedade I, Gunsalus KC, Stoffel M, Rajewsky N (2005) Combinatorial microRNA target predictions. Nat Genet 37(5):495-500 doi:10.1038/ng1536

Krichevsky AM, Gabriely G (2009) miR-21: a small multi-faceted RNA. J Cell Mol Med 13(1):39-53 doi:10.1111/j.1582-4934.2008.00556.x

Krol J, Loedige I, Filipowicz W (2010) The widespread regulation of microRNA biogenesis, function and decay. Nat Rev Genet 11(9):597-610 doi:10.1038/nrg2843

Lewis BP, Burge CB, Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120(1):15-20 doi:10.1016/j.cell.2004.12.035

Lewis NE, Liu X, Li Y, Nagarajan H, Yerganian G, O'Brien E, Bordbar A, Roth AM, Rosenbloom J, Bian C, Xie M, Chen W, Li N, Baycin-Hizal D, Latif H, Forster J, Betenbaugh MJ, Famili I, Xu X, Wang J, Palsson BO (2013) Genomic landscapes of Chinese hamster ovary cell lines as revealed by the *Cricetulus griseus* draft genome. Nat Biotechnol 31(8):759-765 doi:10.1038/nbt.2624

Lin J, Teo S, Lam DH, Jeyaseelan K, Wang S (2012) MicroRNA-10b pleiotropically regulates invasion, angiogenicity and apoptosis of tumor cells resembling mesenchymal subtype of glioblastoma multiforme. Cell Death Dis 3:e398 doi:10.1038/cddis.2012.134

Lin N, Davis A, Bahr S, Borgschulte T, Achtien K, Kayser K (2011) Profiling highly conserved microrna expression in recombinant IgG-producing and parental Chinese hamster ovary cells. Biotechnol Prog 27(4):1163-1171 doi:10.1002/btpr.556

Liu B, Li J, Cairns MJ (2014) Identifying miRNAs, targets and functions. Brief Bioinform 15(1):1-19 doi:10.1093/bib/bbs075

Maccani A, Landes N, Stadlmayr G, Maresch D, Leitner C, Maurer M, Gasser B, Ernst W, Kunert R, Mattanovich D (2014) *Pichia pastoris* secretes recombinant proteins less efficiently than Chinese hamster ovary cells but allows higher space-time yields for less complex proteins. Biotechnol J 9(4):526-537 doi:10.1002/biot.201300305

Maragkakis M, Reczko M, Simossis VA, Alexiou P, Papadopoulos GL, Dalamagas T, Giannopoulos G, Goumas G, Koukis E, Kourtis K, Vergoulis T, Koziris N, Sellis T, Tsanakas P, Hatzigeorgiou AG (2009) DIANA-microT web server: elucidating microRNA functions through target prediction. Nucleic Acids Res 37(Web Server issue):W273-W276 doi:10.1093/nar/gkp292

Maziere P, Enright AJ (2007) Prediction of microRNA targets. Drug Discov Today 12(11-12):452-458 doi:10.1016/j.drudis.2007.04.002

Meents H, Enenkel B, Werner RG, Fussenegger M (2002) $p27^{Kip1}$-mediated controlled proliferation technology increases constitutive sICAM production in CHO-DUKX adapted for growth in suspension and serum-free media. Biotechnol Bioeng 79(6):619-627 doi:10.1002/bit.10322

Miranda KC, Huynh T, Tay Y, Ang YS, Tam WL, Thomson AM, Lim B, Rigoutsos I (2006) A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. Cell 126(6):1203-1217 doi:10.1016/j.cell.2006.07.031

Müller D, Katinger H, Grillari J (2008) MicroRNAs as targets for engineering of CHO cell factories. Trends Biotechnol 26(7):359-365 doi:10.1016/j.tibtech.2008.03.010

Mullokandov G, Baccarini A, Ruzo A, Jayaprakash AD, Tung N, Israelow B, Evans MJ, Sachidanandam R, Brown BD (2012) High-throughput assessment of microRNA activity and function using microRNA sensor and decoy libraries. Nat Methods 9(8):840-846 doi:10.1038/nmeth.2078

Neumüller RA, Betschinger J, Fischer A, Bushati N, Poernbacher I, Mechtler K, Cohen SM, Knoblich JA (2008) Mei-P26 regulates microRNAs and cell growth in the *Drosophila* ovarian stem cell lineage. Nature 454(7201):241-245 doi:10.1038/nature07014

Pasquinelli AE (2012) MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship. Nat Rev Genet 13(4):271-282 doi:10.1038/nrg3162

Pfaffl MW, Horgan GW, Dempfle L (2002) Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res 30(9):e36 doi:10.1093/nar/30.9.e36

Rehmsmeier M, Steffen P, Hochsmann M, Giegerich R (2004) Fast and effective prediction of microRNA/target duplexes. RNA 10(10):1507-1517 doi:10.1261/rna.5248604

Reiner A, Yekutieli D, Benjamini Y (2003) Identifying differentially expressed genes using false discovery rate controlling procedures. Bioinformatics 19(3):368-375 doi:10.1093/bioinformatics/btf877

Ritchie W, Flamant S, Rasko JE (2009) Predicting microRNA targets and functions: traps for the unwary. Nat Methods 6(6):397-398 doi:10.1038/nmeth0609-397

Shin C, Nam JW, Farh KK, Chiang HR, Shkumatava A, Bartel DP (2010) Expanding the microRNA targeting code: functional sites with centered pairing. Mol Cell 38(6):789-802 doi:10.1016/j.molcel.2010.06.005

Strotbek M, Florin L, Koenitzer J, Tolstrup A, Kaufmann H, Hausser A, Olayioye MA (2013) Stable microRNA expression enhances therapeutic antibody productivity of Chinese hamster ovary cells. Metab Eng 20:157-166 doi:10.1016/j.ymben.2013.10.005

Sturn A, Quackenbush J, Trajanoski Z (2002) Genesis: cluster analysis of microarray data. Bioinformatics 18(1):207-208 doi:10.1093/bioinformatics/18.1.207

Vella MC, Choi EY, Lin SY, Reinert K, Slack FJ (2004) The *C. elegans* microRNA *let-7* binds to imperfect *let-7* complementary sites from the *lin-41* 3'UTR. Genes Dev 18(2):132-137 doi:10.1101/gad.1165404

Vishwanathan N, Le H, Jacob NM, Tsao YS, Ng SW, Loo B, Liu Z, Kantardjieff A, Hu WS (2014) Transcriptome dynamics of transgene amplification in Chinese hamster ovary cells. Biotechnol Bioeng 111(3):518-528 doi:10.1002/bit.25117

Wang X (2008) miRDB: a microRNA target prediction and functional annotation database with a wiki interface. RNA 14(6):1012-1017 doi:10.1261/rna.965408

Xu X, Nagarajan H, Lewis NE, Pan S, Cai Z, Liu X, Chen W, Xie M, Wang W, Hammond S, Andersen MR, Neff N, Passarelli B, Koh W, Fan HC, Wang J, Gui Y, Lee KH, Betenbaugh MJ, Quake SR, Famili I, Palsson BO, Wang J (2011) The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. Nat Biotechnol 29(8):735-741 doi:10.1038/nbt.1932

Identification of microRNAs specific for high producer CHO cell lines using steady-state cultivation

Andreas Maccani[1,2], Matthias Hackl[2], Christian Leitner[3], Willibald Steinfellner[2], Alexandra B. Graf[1,4], Nadine E. Tatto[1], Michael Karbiener[5], Marcel Scheideler[5], Johannes Grillari[2], Diethard Mattanovich[1,2], Renate Kunert[1,2], Nicole Borth[1,2], Reingard Grabherr[2], Wolfgang Ernst[1,2]

[1] Austrian Centre of Industrial Biotechnology (ACIB GmbH), Muthgasse 11, 1190 Vienna, Austria

[2] Department of Biotechnology, VIBT – BOKU – University of Natural Resources and Life Sciences Vienna, Muthgasse 18, 1190 Vienna, Austria

[3] Department of Food Science and Technology, VIBT – BOKU – University of Natural Resources and Life Sciences Vienna, Muthgasse 18, 1190 Vienna, Austria

[4] School of Bioengineering, University of Applied Sciences FH-Campus Vienna, Muthgasse 56, 1190 Vienna, Austria

[5] Institute for Genomics and Bioinformatics, Graz University of Technology, Petersgasse 14, 8010 Graz, Austria

Correspondence: Andreas Maccani, ACIB GmbH, Muthgasse 11, 1190 Vienna, Austria
Email: andreas.maccani@acib.at

Table S1: Primers used for qRT-PCR

Target	Primer sequence (5'–3') or Qiagen miScript Primer Assay
Actr5 sense	CTCCTTCCAGGTTCAGCTTG
Actr5 antisense	GGCACAATGTTCCTTGAGGT
Gapdh sense	GTAAGAAGCCCACCCTGGA
Gapdh antisense	GTGAGGGAGATGATCGGTGT
3D6scFv-Fc sense	CCCAAGCTGCTGATCTACAA
3D6scFv-Fc antisense	GATGGTCAGGGTGAACTCG
HSA sense	CCTGGAAGTGGACGAGACATAC
HSA antisense	GTCTGCTTCTTGATCTGCCTTT
let-7b-5p	MS00001225
let-7c-5p	MS00005852
miR-100-5p	MS00032214
miR-10b-5p	MS00032249
miR-125b-5p	MS00005992
miR-19a-3p	MS00001302
miR-185-5p	MS00001736
miR-193a-3p	MS00001785
miR-21-5p	MS00011487
miR-221-3p	MS00032585
miR-350-3p	MS00007938
miR-99a-5p	MS00033117

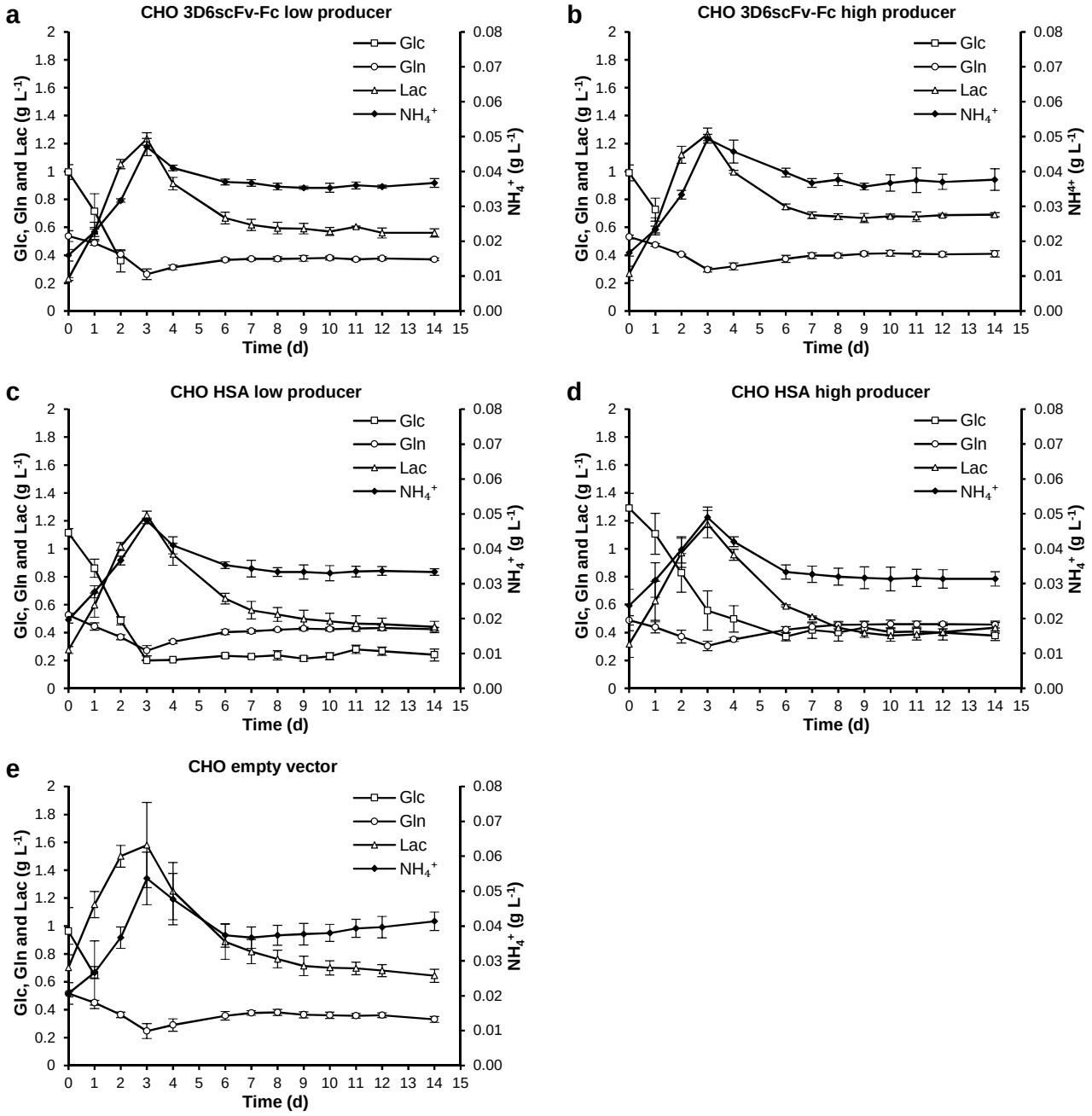


Fig. S1 Time courses of steady-state cultivations. Glucose (Glc), glutamine (Gln), lactate (Lac) and ammonium (NH_4^+) concentration of (**a**) CHO 3D6scFv-Fc low-producer, (**b**) CHO 3D6scFv-Fc high-producer, (**c**) CHO HSA low-producer, (**d**) CHO HSA high-producer and (**e**) CHO empty vector (non-producer). Cells were cultivated in a 0.8 L cell culture bioreactor. After three days of batch cultivation, the process was switched to continuous cultivation (dilution rate D = 0.5 d^{-1}). The culture volume was maintained at a constant level of 400 mL. Data represent mean values of three independent cultivations (error bars: SD). No data points are shown for glucose concentration below the detection limit of 0.2 g L^{-1}

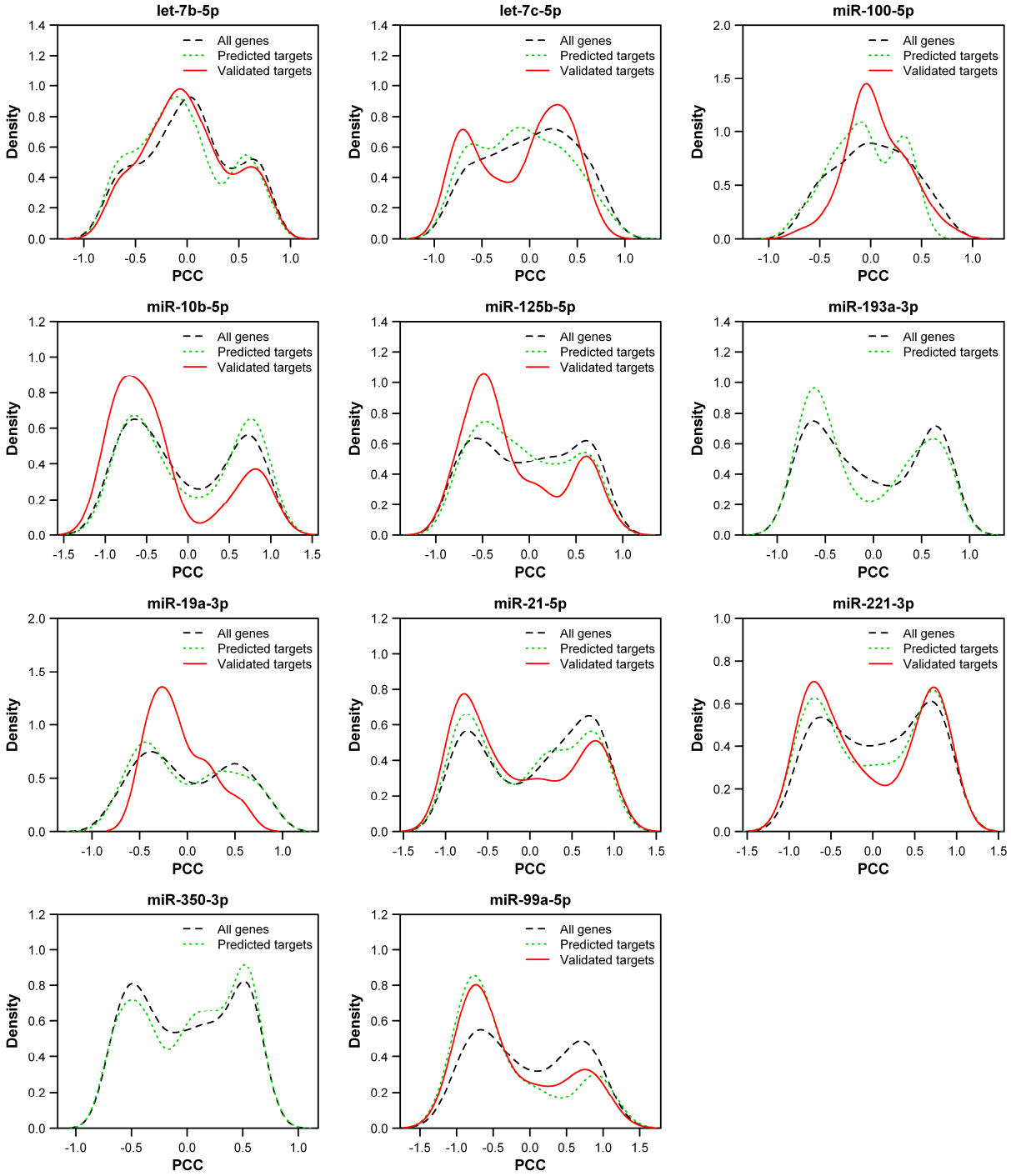


Fig. S2 Distribution of Pearson correlation coefficients (PCC) between miRNAs and validated or predicted targets. Superimposed kernel density plots were computed with equal bandwidths. All genes: 2843 mRNAs which were identified as differentially expressed (adj. $p < 0.05$ and fold change > 1.5) between the cell lines used in this study. Predicted targets: Differentially expressed target mRNAs that were computationally predicted by more than half of nine applied algorithms. Validated targets: Differentially expressed target mRNAs which have been experimentally validated in human, mouse or rat (miRTarBase 4.5)

4. Results and discussion

4.1 CHO cells versus *Pichia pastoris*: A system comparison

Chinese hamster ovary (CHO) cells are currently the workhorses of the biopharmaceutical industry. However, various eukaryotic hosts ranging from microbes such as yeasts and filamentous fungi to higher eukaryotes including plant, insect as well as other mammalian cell lines have been utilized for recombinant protein production, where each expression system has specific benefits and drawbacks. Yeasts such as *Pichia pastoris* are considered as promising alternatives, as they can grow more rapidly to much higher cell densities than mammalian cells. On the other hand, yeasts cannot compete against the high secretory capacity of mammalian cells. However, although a lot of valuable information about the capability of CHO cells and *P. pastoris* for heterologous protein production is available, a quantitative comparison of process relevant data is very difficult. The specific product secretion rate (q_P) strongly depends on the particular protein, however, the proteins produced in microbial systems are in general different from those expressed in mammalian cells. Therefore, we analyzed process relevant parameters of high producing *P. pastoris* strains and CHO cell lines secreting the same two model proteins in comparable fed batch processes (Publication B).

Human serum albumin (HSA) and the 3D6 anti-HIV-1 single-chain Fv-Fc fusion antibody (3D6scFv-Fc), two proteins with very different structure and complexity, were selected as model proteins in order to challenge the expression systems in different ways. HSA is a monomeric protein of 67 kDa composed of three domains with a high content of α-helices (67%) but no β-sheets and 17 intramolecular disulfide bonds (Curry et al. 1998), and it can be expressed at very high levels in *P. pastoris* (Kobayashi et al. 2000). The more complex homodimeric protein 3D6scFv-Fc (Figure 6) has about 110 kDa and contains the Fc-specific glycosylation site of an IgG1 antibody.

Stable high producing CHO cell lines and *P. pastoris* strains were established by increasing the transgene copy number via gene amplification and cultivated in standard fed batch processes using the same bioreactor system. In the following, the most important findings of Publication B as well as additional aspects of this study are summarized and discussed.

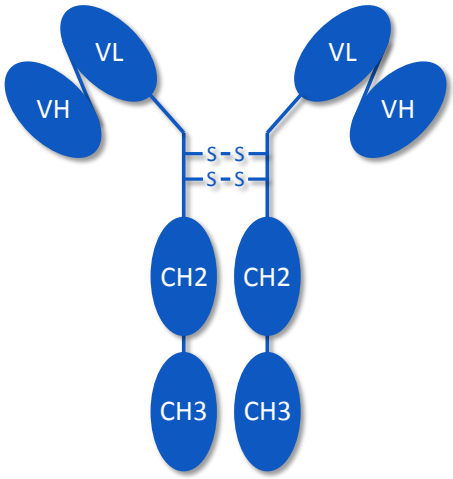


Figure 6. Schematic structure of a single-chain Fv-Fc fusion antibody. The variable heavy chain (VH) and the variable light chain (VH) domain are combined by a $(GGGGS)_3$ linker and fused via the hinge region to the IgG1 fragment crystallizable (Fc) region, which is composed of two constant domains (CH2 and CH3).

4.1.1 Cell growth and cell density

One major limitation of CHO cells is the low specific growth rate. In the present study, the maximal specific growth rate (μ_{max}) was about 5–6-fold lower compared to *P. pastoris*. Consequently more time is required to build up biomass which leads to significantly longer process times. To compare the biomass density of both expression systems, the cell dry mass (CDM) concentration was used instead of the cell concentration which is typically used in cell culture technology. The maximal CDM concentrations were about 140–300-fold lower for CHO cells compared to *P. pastoris*. In the applied standard (non-optimized) fed batch process, the maximal CDM concentration of the CHO 3D6scFv-Fc producer was 0.604 g L^{-1} which corresponds to a viable cell concentration of 2.26×10^9 cells L^{-1}. However, optimized fed batch processes (media composition, cultivation strategy) can reach cell densities of $10–15\times10^9$ cells L^{-1} today (De Jesus and Wurm 2011). Even higher cell densities ($>20\times10^9$ cells L^{-1}) can be achieved using perfusion. Clincke et al. (2013) even reported a CHO cell density of $>200\times10^9$ cells L^{-1} in a perfusion process using a wave-induced bioreactor. But perfusion processes are technically more challenging because spent medium needs to be continuously separated from the cells and removed as well as fresh medium added.

4.1.2 Productivity

The most significant differences could be observed comparing the biomass-specific product secretion rates q_P ($mg_{product}\ g_{CDM}^{-1}\ h^{-1}$). In *P. pastoris*, q_P was strongly product dependent (40 times lower for 3D6scFv-Fc compared to HSA), whereas in CHO cells q_P was very similar for

both model proteins. Comparing both organisms, the mean q_P of the CHO cell lines was >1000-fold higher for 3D6scFv-Fc and 26-fold higher for HSA. These results clearly demonstrates that CHO cells are superior compared to *P. pastoris* with respect to recombinant protein secretion.

The space-time yield (STY) is an important criterion to assess the economic efficiency of the fermentation process. The STY depends on the available viable biomass, the q_P and the process duration. Because of the higher biomass densities and the shorter process times, *P. pastoris* can partly compensate for the lower secretion potential. Thus, the STY of the HSA producer was 9.2-fold higher compared to CHO cells. However, due to the very low q_P of the 3D6scFv-Fc producing strain, the STY was 9.6-fold lower for *P. pastoris*. Consequently, it may be speculated that antibody and antibody fragments can be produced more economically in CHO cells, while *P. pastoris* is the better choice for less complex proteins such as HSA.

4.1.3 Product quality

Beside high productivity, an expression system should deliver products of high quality which are biological active. This includes correct protein folding and assembly as well as proper glycosylation. Whereas CHO cells secreted only products of correct size in our study, little amounts of degraded products were detected by Western blot analysis in the *P. pastoris* processes although a protease deficient strain (SMD1168H) was used (Figure 2, Publication B). Proteolytic degradation has an effect on process efficiency as well as product purity and is still a major problem in yeasts, although various methods to reduce proteolytic degradation including the generation of protease deficient strains have been established (Macauley-Patrick et al. 2005; Sinha et al. 2005; Wu et al. 2013). One major advantage of *P. pastoris* is the low amount of host cell proteins (HCPs) in the culture supernatant (Mattanovich et al. 2009). In CHO cells, we observed an increasing amount of HCPs in the supernatant in the course of the fed batch (Figure 7). The results of Tait et al. (2012) indicated that HCPs in the supernatant mainly arise through cell lysis rather than by secretion. Hence, it is crucial to maintain a high viability until the end of the process in order to minimize HCP contaminations. However, in our cultivations the product was clearly the most abundant protein in the supernatant indicating a high relative purity.

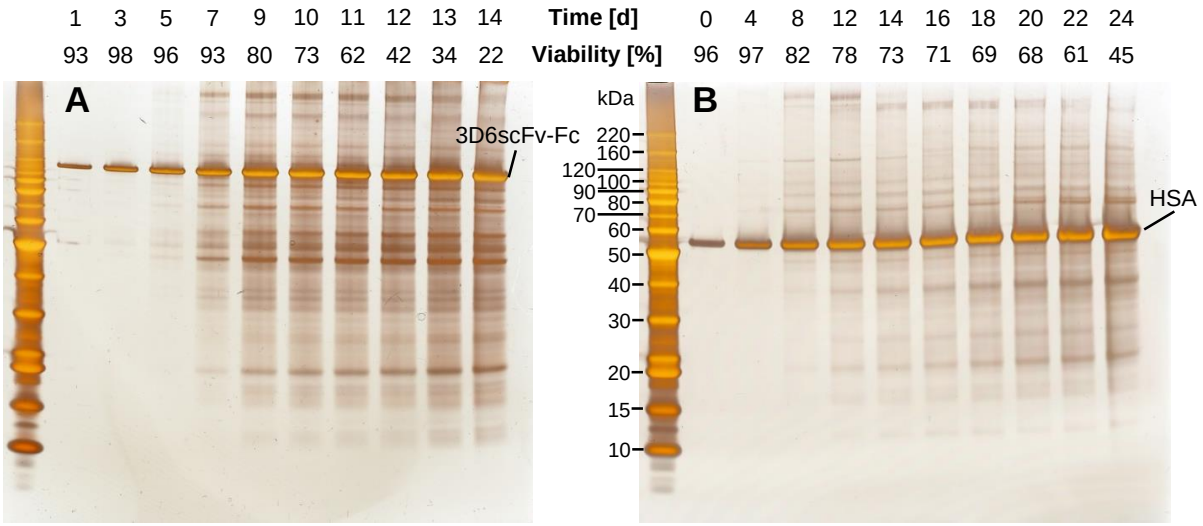


Figure 7. Total protein content in the supernatant during fed batch cultivation. Silver stained SDS-PAGE under non-reducing conditions of the supernatants of (**A**) CHO 3D6scFv-Fc producer and (**B**) CHO HSA producer. Equal volumes of supernatant were loaded onto the gel. The results of one representative fed batch cultivation are shown.

As expected, differences were observed in the N-glycosylation pattern of the 3D6scFv-Fc antibody. CHO cells produced mainly fucosylated Fc glycans of the complex-type biantennary structure which are also typically found in human IgGs (Zauner et al. 2013). *P. pastoris*-derived 3D6scFv-Fc antibodies contained exclusively high-mannose-type N-glycans (Man9–Man16) or were partly unglycosylated. The differences in Fc glycosylation did not impair epitope binding, however, it is known that effector functions such as antibody-dependent cellular cytotoxicity (ADCC) and complement-dependent cytotoxicity (CDC) are strongly affected by the glycan structure (Jefferis 2009). Hence, CHO cells can produce fully functional IgGs or scFv-Fc antibodies by nature, whereas glycoengineered strains would be required for *P. pastoris* (Ye et al. 2011).

4.1.4 Cross-species comparison to improve recombinant protein production

The results of this study clearly show that the protein secretion machinery is much more efficient in CHO cells than in *P. pastoris*. Whereas CHO cells produced both model proteins at the same level, 3D6scFv-Fc secretion was significantly impaired in *P. pastoris*. The mechanisms that control and limit recombinant protein formation and secretion in both organisms are still poorly understood. Hence, a comparative genome-wide analysis of both eukaryotic expression systems might be very useful to identify conserved mechanisms and pathways that influence heterologous protein synthesis and secretion as well as cell growth.

Conserved genes and pathways that are differentially regulated between both organisms and/or both model proteins may be crucial for efficient protein production or high cell growth, and therefore potential targets for cell engineering.

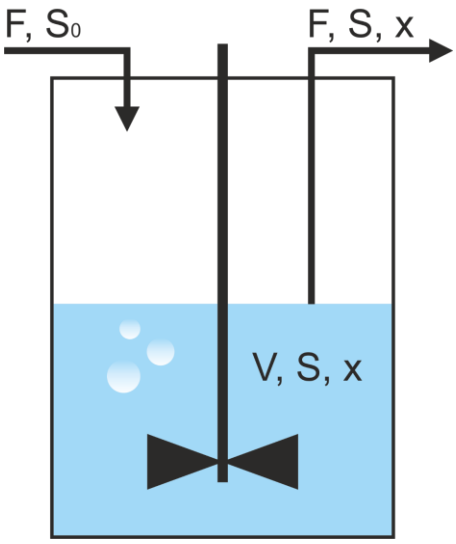
4.2 Global mRNA expression profiling in recombinant CHO cell lines

In this study, the mRNA expression profiles of high, low and non-producing recombinant CHO cell lines were analyzed and compared in order to identify genes and pathways that are involved in heterologous protein synthesis and secretion. For this purpose, CHO cell lines expressing 3D6scFv-Fc and HSA were generated and low and high producers were selected at different stages of transgene amplification (Publication B and C). In addition a mock transfected CHO cell lines was established (non-producer). All five cell lines were cultivated in biological triplicates in 800 mL DASGIP bioreactors under steady-state conditions using a chemostat process (Publication C).

4.2.1 Comparable conditions by steady-state cultivation

A prerequisite for the comparison of expression profiles of different cell lines is that the samples are generated under comparable and defined condition in order to obtain reproducible and meaningful data. In previously published CHO omics studies, cell samples were generally taken from the exponential growth phase of batch cultivations. However, in simple batch cultures the physico-chemical conditions are very dynamic due to nutrient consumption and accumulation of metabolites, and have a significant impact on the cell's transcriptome (Hernandez Bort et al. 2012). Hence, the time of sampling is crucial for the outcome of comparative gene expression studies when batch cultures are used. Considering the sometimes considerable differences in cell growth of different recombinant cell lines as well as host cell lines, makes it clear that the optimal time of sampling is very difficult to define. Consequently the chemostat (from *chem*ical environment is *stat*ic) which provides a defined and constant environment is the ideal cultivation mode for gene expression experiments comparing different cell lines (Hoskisson and Hobbs 2005).

In a chemostat fresh medium is continuously added, while culture broth is continuously removed at the same flow rate (F) in order to keep the culture volume (V) constant (Figure 8). The cell concentration (x) levels off at a certain value where growth is limited by the limiting nutrient (glucose). This enables the cultivation of cells in a physiological steady-state, where the cells grow at a constant rate and all culture parameters (e.g. nutrient, metabolite and product concentrations) remain constant. In steady-state the specific growth rate (μ) of the cells is equal to the dilution rate (D), where D is defined as the flow rate (F) divided by the culture volume (V). Consequently, cells can be grown at a defined and constant specific growth rate (μ), and μ can be set to any value lower than μ_{max} by adjusting the flow rate (F).



$$V = \text{const.}$$

$$D = \frac{V}{F}$$

$$\frac{dx}{dt} = \mu x - Dx$$

Steady-state:

$$\frac{dx}{dt} = \mu x - Dx = 0 \quad \Rightarrow \quad \mu = D$$

Figure 8. Principle of steady-state cultivation using a chemostat. Flow rate (F), substrate concentration (S), cell concentration (x), culture volume (V), dilution rate (D) and specific growth rate (μ)

Chemostat cultivation of CHO cells to achieve steady-state conditions was already described more than two decades ago (Hayter et al. 1992). However, it was not applied in recent CHO-omics studies, whereas the chemostat is frequently used for yeasts (Baumann et al. 2011; Dragosits et al. 2009).

4.2.2 A novel CHO-specific microarray

The publication of the first genome sequence of a CHO cell line (CHO-K1) in 2011 was also a major milestone for transcriptome analysis of CHO cells (Xu et al. 2011b). Due to the lack of publicly available CHO sequence information before 2011, our microarray analyses of CHO cells relied on the cross-species hybridization to commercially available microarrays designed for closely related species such as mouse and rat (Ernst et al. 2006; Hernandez Bort et al. 2012; Trummer et al. 2008). However, although CHO transcripts show a high degree of sequence similarity (92%) to their corresponding mouse and rat homologs (Ernst et al. 2006), cross-species hybridization is accompanied by a lower sensitivity.

Based on the publicly available genomic sequence data of the CHO-K1 cell lines, we designed 60-mer oligonucleotides probes for 20,650 unique transcripts, which were spotted in duplicates using the 4x44k oligonucleotide microarray design offered by Agilent (CA, USA). In this experiment, we analyzed five different cell lines in biological triplicates and technical duplicates (dye-swap) and 11,255 unique transcripts could be reliably detected in at least one of the five cell lines (based on Agilent GeneSpring 12.6 analysis).

The results of the microarray analysis were validated using quantitative reverse transcriptase PCR (qRT-PCR). For this purpose 38 mRNAs and two housekeeping genes (*Gapdh* and *Actr5*) were selected. *Gapdh* is highly and *Actr5* moderately expressed in CHO cells and the expression levels were very stable between the different samples. To distinguish between mature mRNA and genomic DNA, the primers were preferably designed at exon-exon junction or that the amplicon spans at least one intron. qRT-PCRs were conducted for all 15 samples (five cell lines, three biological replicates) in technical duplicates (independent qRT-PCR runs). For statistical analysis and determination of the ratios using two housekeeping genes, the software REST 2009 was used. The primer pairs for seven mRNAs generated wrong or additional products or primer dimers and so these results were not included in the analysis. However, the expression ratios of 30 of the remaining 31 genes corresponded very well with the data obtained by the microarray analysis (Figure 9 – Figure 14). Only for one gene (*Akr1b7*), the expression data of qRT-PCR and microarray analysis were significantly different (Figure 9). One possible explanation would be an alternative splicing variant. The 60-mer oligonucleotide probe on the microarray is specific for the end of the 3' untranslated region (UTR), whereas the qRT-PCR primers bind to exon 1 and exon 2, respectively. Hence, a potential splicing variant lacking exon 1 or exon 2 is not detected by qRT-PCR. On the other hand, alternative splicing that affects the 3' UTR might result in a transcript that escapes microarray detection.

Nevertheless, the results of the qRT-PCR analysis confirmed that our CHO-specific microarray provided accurate and reliable data.

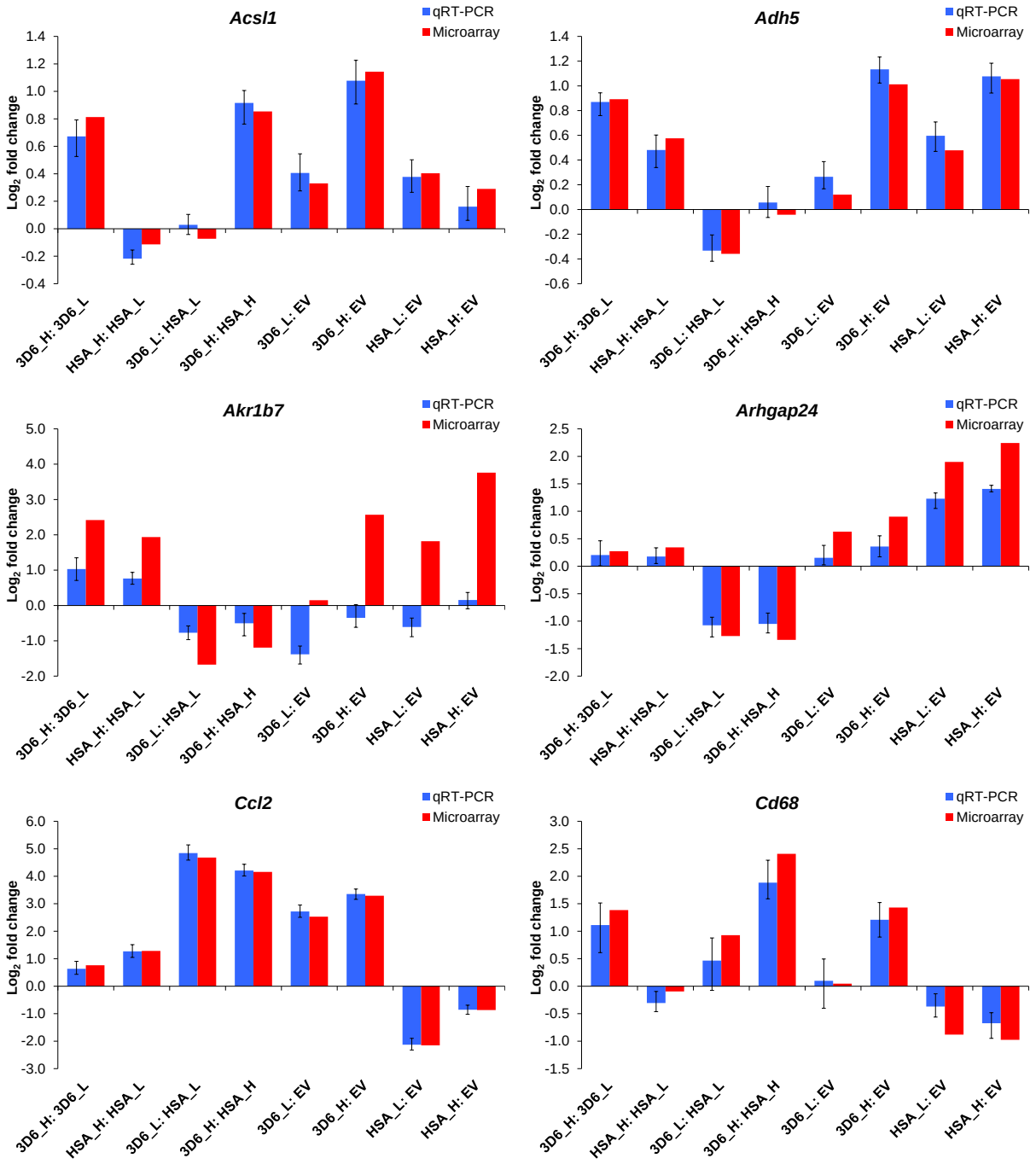


Figure 9. Comparison of gene expression determined by qRT-PCR and microarray (*Acsl1–Cd68*). 3D6_H, CHO 3D6scFv-Fc high producer; 3D6_L, CHO 3D6scFv-Fc low producer; HSA_H, CHO HSA high producer; HSA_L, CHO HSA low producer; EV, CHO empty vector (non-producer). qRT-PCR data were normalized using two endogenous controls (*Gapdh* and *Actr5*). Data represent mean values of samples obtained from three independent steady-state cultivations (n = 3, error bars: SE)

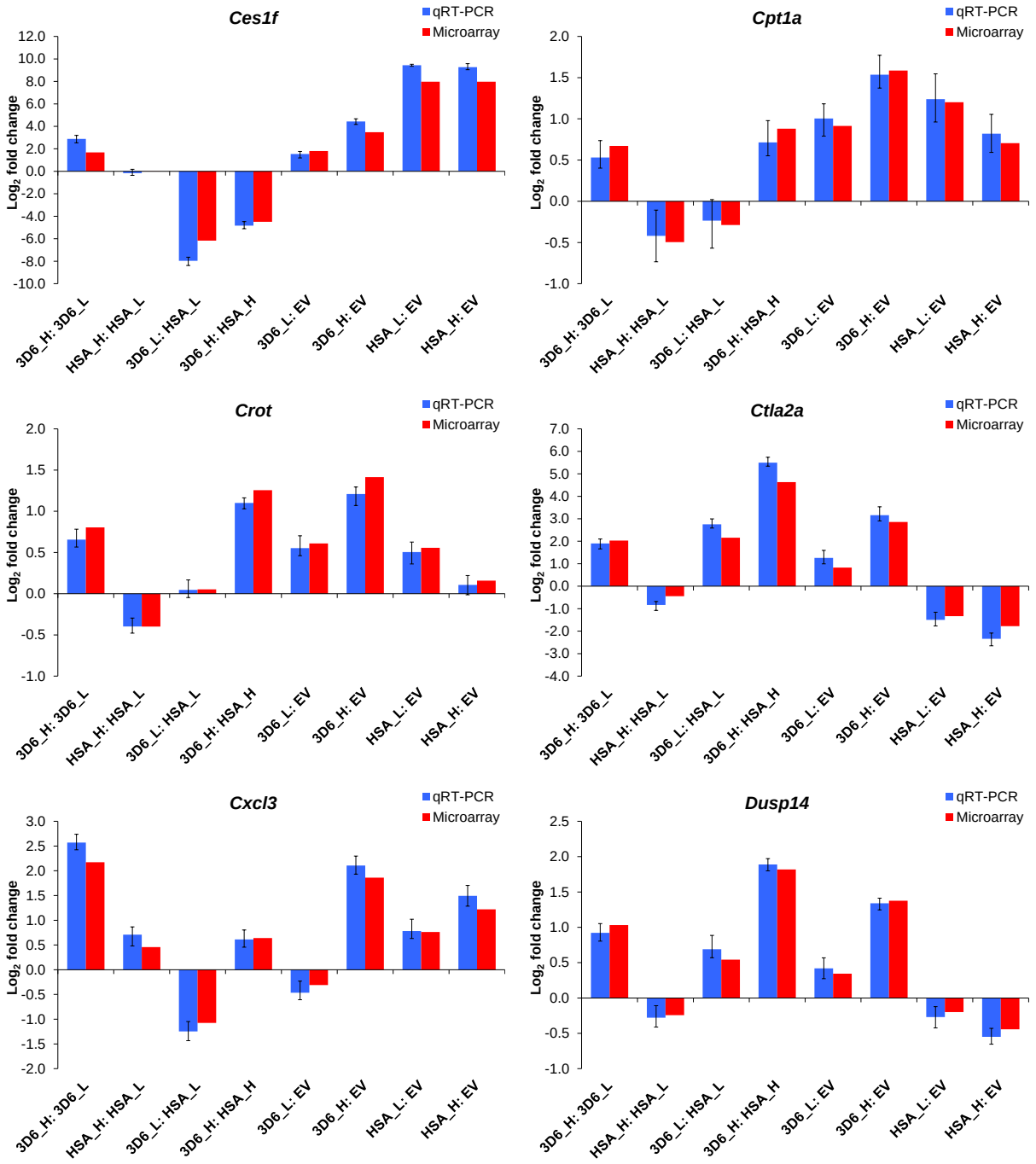


Figure 10. Comparison of gene expression determined by qRT-PCR and microarray (*Cesf1f–Dusp14*). 3D6_H, CHO 3D6scFv-Fc high producer; 3D6_L, CHO 3D6scFv-Fc low producer; HSA_H, CHO HSA high producer; HSA_L, CHO HSA low producer; EV, CHO empty vector (non-producer). qRT-PCR data were normalized using two endogenous controls (*Gapdh* and *Actr5*). Data represent mean values of samples obtained from three independent steady-state cultivations (n = 3, error bars: SE)

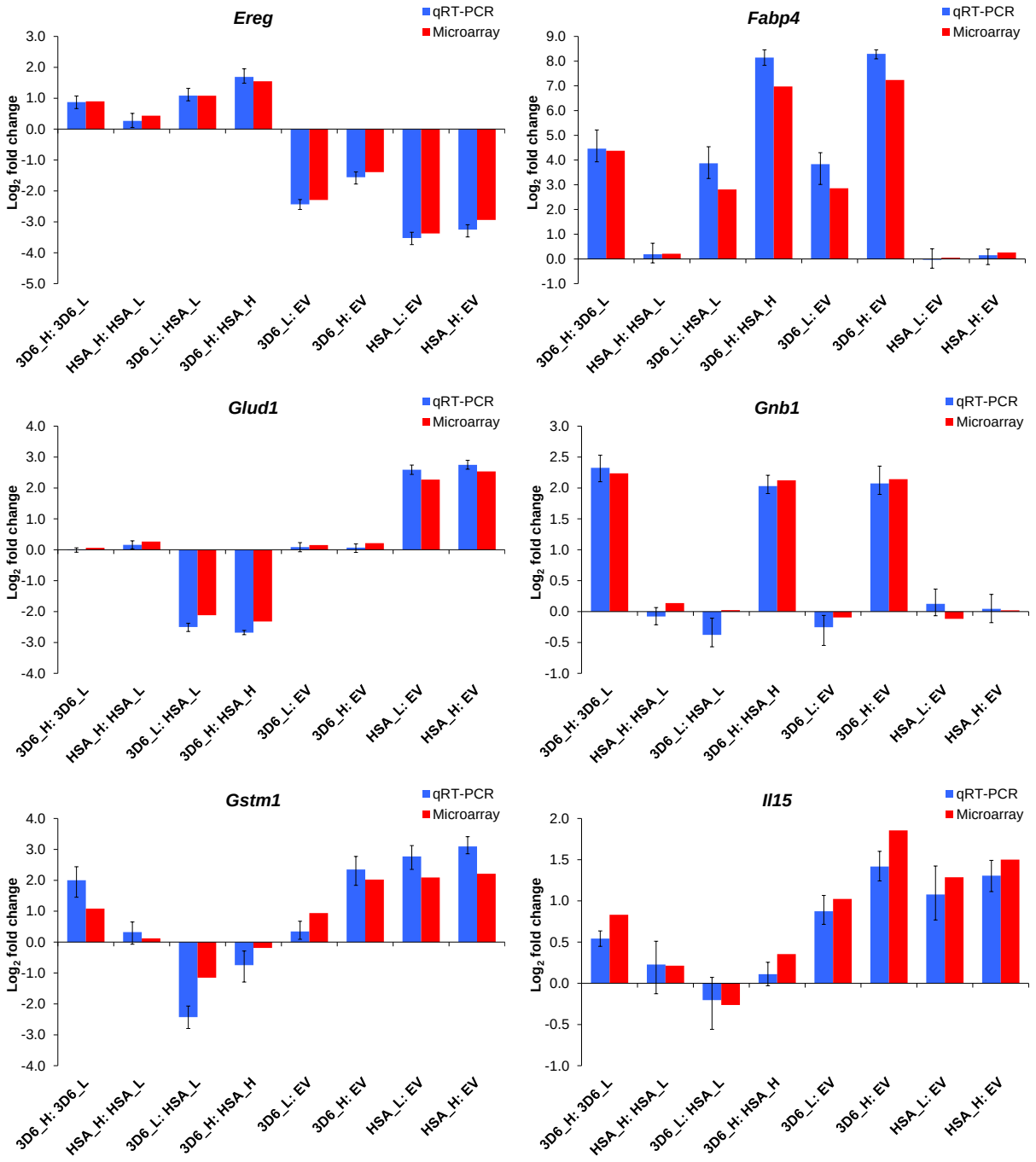


Figure 11. Comparison of gene expression determined by qRT-PCR and microarray (*Ereg–Il15*). 3D6_H, CHO 3D6scFv-Fc high producer; 3D6_L, CHO 3D6scFv-Fc low producer; HSA_H, CHO HSA high producer; HSA_L, CHO HSA low producer; EV, CHO empty vector (non-producer). qRT-PCR data were normalized using two endogenous controls (*Gapdh* and *Actr5*). Data represent mean values of samples obtained from three independent steady-state cultivations (n = 3, error bars: SE)

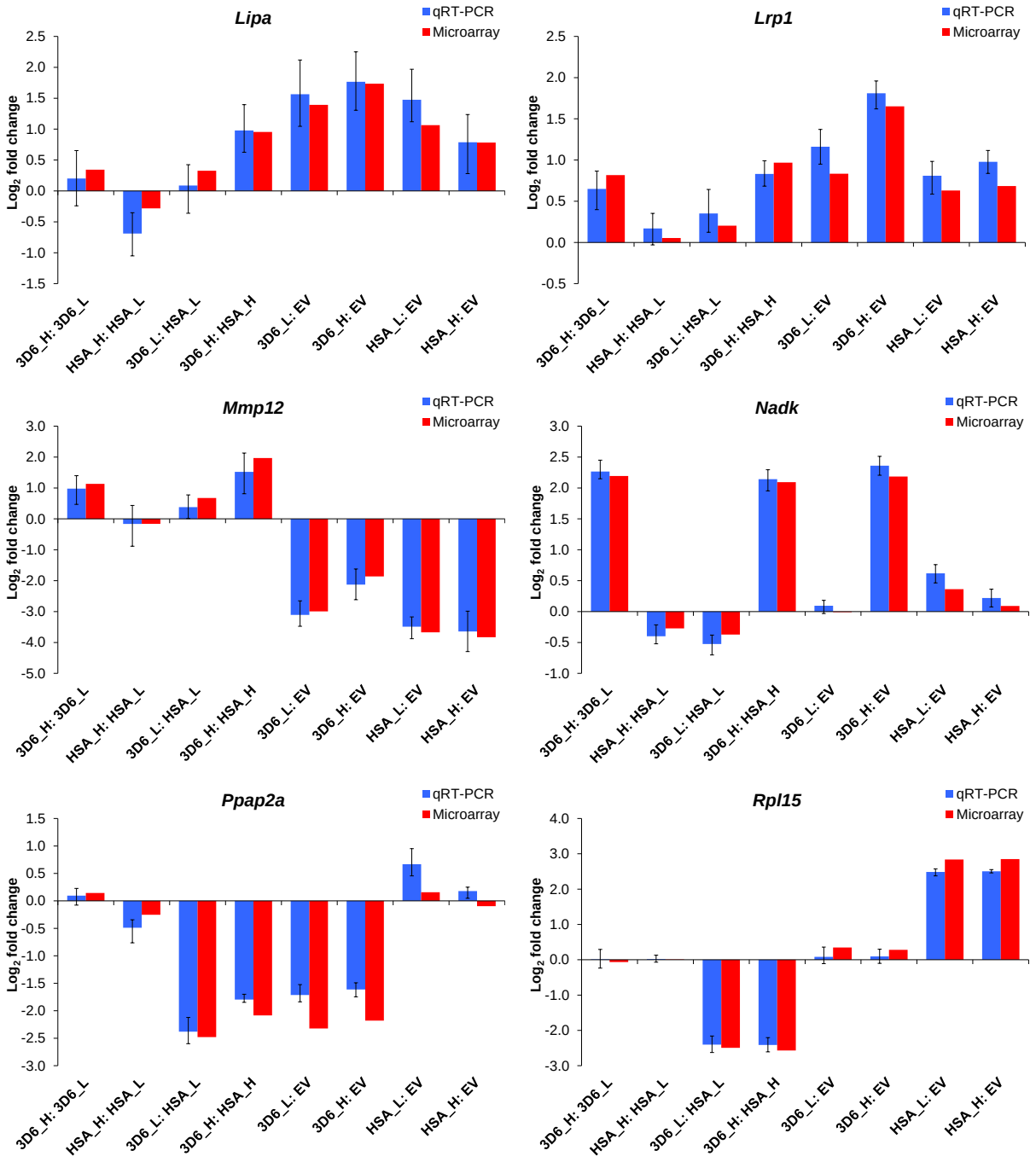


Figure 12. Comparison of gene expression determined by qRT-PCR and microarray (*Lipa–Rpl15*). 3D6_H, CHO 3D6scFv-Fc high producer; 3D6_L, CHO 3D6scFv-Fc low producer; HSA_H, CHO HSA high producer; HSA_L, CHO HSA low producer; EV, CHO empty vector (non-producer). qRT-PCR data were normalized using two endogenous controls (*Gapdh* and *Actr5*). Data represent mean values of samples obtained from three independent steady-state cultivations (n = 3, error bars: SE)

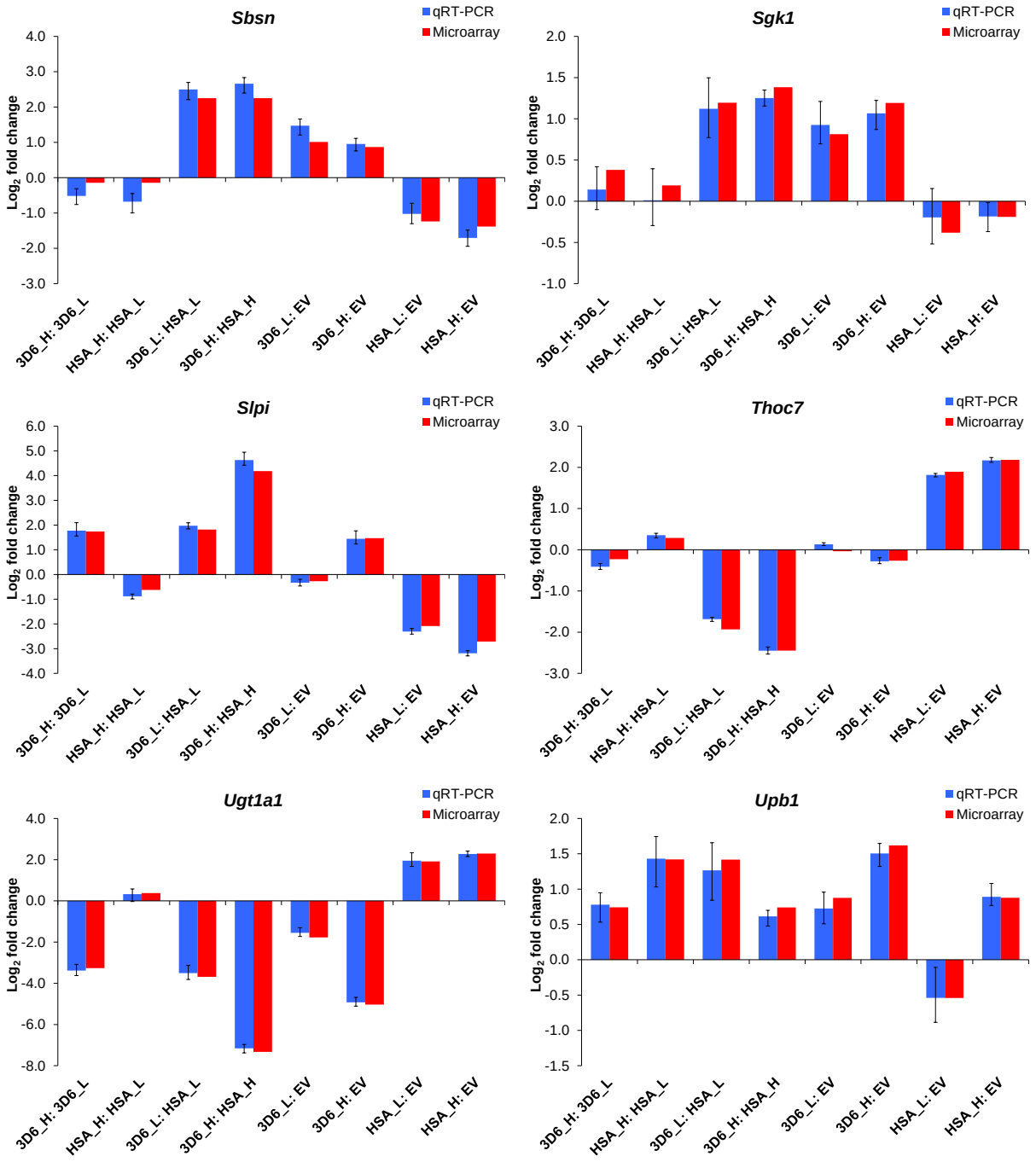


Figure 13. Comparison of gene expression determined by qRT-PCR and microarray (*Sbsn–Upb1*). 3D6_H, CHO 3D6scFv-Fc high producer; 3D6_L, CHO 3D6scFv-Fc low producer; HSA_H, CHO HSA high producer; HSA_L, CHO HSA low producer; EV, CHO empty vector (non-producer). qRT-PCR data were normalized using two endogenous controls (*Gapdh* and *Actr5*). Data represent mean values of samples obtained from three independent steady-state cultivations (n = 3, error bars: SE)

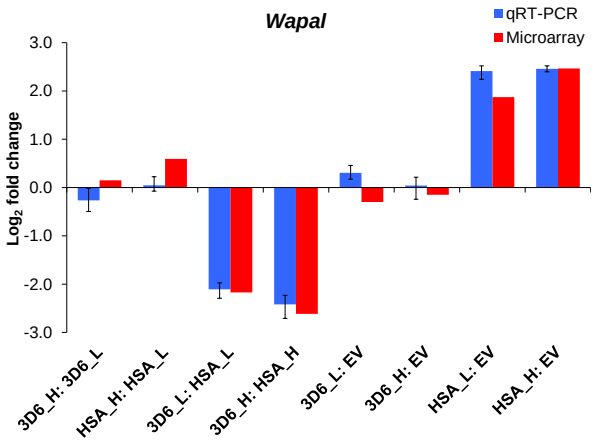


Figure 14. Comparison of gene expression determined by qRT-PCR and microarray (*Wapal*). 3D6_H, CHO 3D6scFv-Fc high producer; 3D6_L, CHO 3D6scFv-Fc low producer; HSA_H, CHO HSA high producer; HSA_L, CHO HSA low producer; EV, CHO empty vector (non-producer). qRT-PCR data were normalized using two endogenous controls (*Gapdh* and *Actr5*). Data represent mean values of samples obtained from three independent steady-state cultivations (n = 3, error bars: SE)

4.2.3 Product/clone dependency of mRNA expression

In order to identify differences in the gene expression profiles of the five CHO cell lines analyzed in this study, following eight comparisons were made.

- CHO 3D6scFv-Fc high producer vs. CHO 3D6scFv-Fc low producer (3D6_H: 3D6_L)
- CHO HSA high producer vs. CHO HSA low producer (HSA_H: HSA_L)
- CHO 3D6scFv-Fc low producer vs. CHO HSA low producer (3D6_L: HSA_L)
- CHO 3D6scFv-Fc high producer vs. CHO HSA high producer (3D6_H: HSA_H)
- CHO 3D6scFv-Fc low producer vs. CHO empty vector (3D6_L: EV)
- CHO 3D6scFv-Fc high producer vs. CHO empty vector (3D6_H: EV)
- CHO HSA low producer vs. CHO empty vector (HSA_L: EV)
- CHO HSA high producer vs. CHO empty vector (HSA_H: EV)

Genes with a fold change (FC) ≥ 1.5 and an adjusted *p*-value (adj. *p*) ≤ 0.05 (multiple testing correction based on Benjamini and Yekutieli) were considered as differentially expressed. A total of 2842 genes were differentially expressed in at least one of the eight comparisons. The specific product secretion rates (q_P) were ~8-fold higher in the high producer cell lines compared to the low producers for both model proteins. However, significantly more genes were differentially expressed between high and low producer of 3D6scFv-Fc than between

high and low producer of HSA (Table 1). On the other hand, a higher number of genes was differentially expressed between the HSA producers and the non-producer (CHO empty vector) than between the 3D6scFv-Fc producers and the non-producer. Most differentially expressed genes were, however, identified between CHO 3D6scFv-Fc high producer and CHO HSA high producer.

Table 1. Number of differentially expressed genes (Fold change ≥ 1.5 and adj. $p \leq 0.05$)

	3D6_H: 3D6_L	**HSA_H: HSA_L**	**3D6_L: HSA_L**	**3D6_H: HSA_H**	**3D6_L: EV**	**3D6_H: EV**	**HSA_L: EV**	**HSA_H: EV**
Up	362	204	614	773	262	517	504	673
Down	339	84	377	734	222	467	697	817
Total	701	288	991	1507	484	984	1201	1490

Venn diagrams were used to compare and visualize overlaps between the different datasets (Figure 15). Only few genes (16.3% of HSA_H: HSA_L) were commonly up- or down-regulated between high and low producers for both model proteins (Figure 15A). Comparing 3D6scFv-Fc low and high producer producers against the non-producer showed that 58.1% of the up- or down-regulated genes in low producer were also up- or down-regulated in the high producer (Figure 15B). For HSA, the overlap was higher with 76.6% (Figure 15C). 49.6% of the up- or down-regulated genes between 3D6scFv-Fc low producer and non-producer were also up- or down-regulated in the HSA low producer (Figure 15D). Regarding the high producers, 39.5% of the up- or down-regulated genes between 3D6scFv-Fc high producer and non-producer were also up- or down-regulated in the HSA high producer (Figure 15E). Comparing all producers against the non-producer showed that 60 genes were commonly up-regulated and 85 commonly down-regulated (Figure 15F).

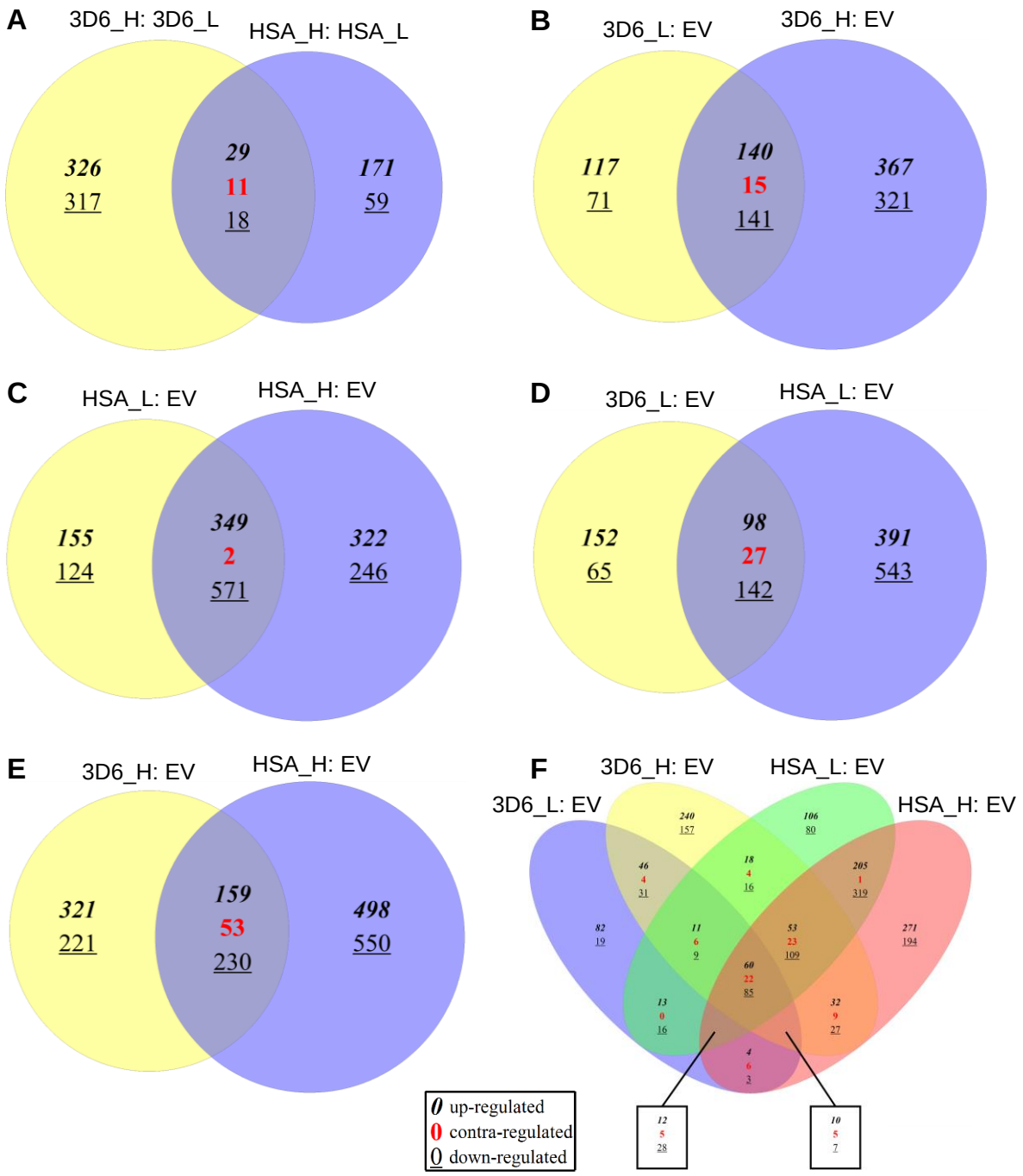


Figure 15. Number of commonly and exclusively differentially expressed genes. 3D6_H, CHO 3D6scFv-Fc high producer; 3D6_L, CHO 3D6scFv-Fc low producer; HSA_H, CHO HSA high producer; HSA_L, CHO HSA low producer; EV, CHO empty vector (non-producer). Venn diagrams were generated using VennPlex (Cai et al. 2013).

In summary, we observed a considerable divergence in the affected genes that are potentially related with recombinant protein biosynthesis and secretion between cell lines producing different proteins. The differences in gene expression were more distinct between 3D6scFv-Fc and HSA producers than between high and low producers. This was also observed by

Principal component analysis (PCA) using the whole dataset (including all genes that were not significantly differentially expressed), where component 1 (50.4%) clearly separates 3D6scFv-Fc producers and HSA producers, but not high and low producers (Figure 16).

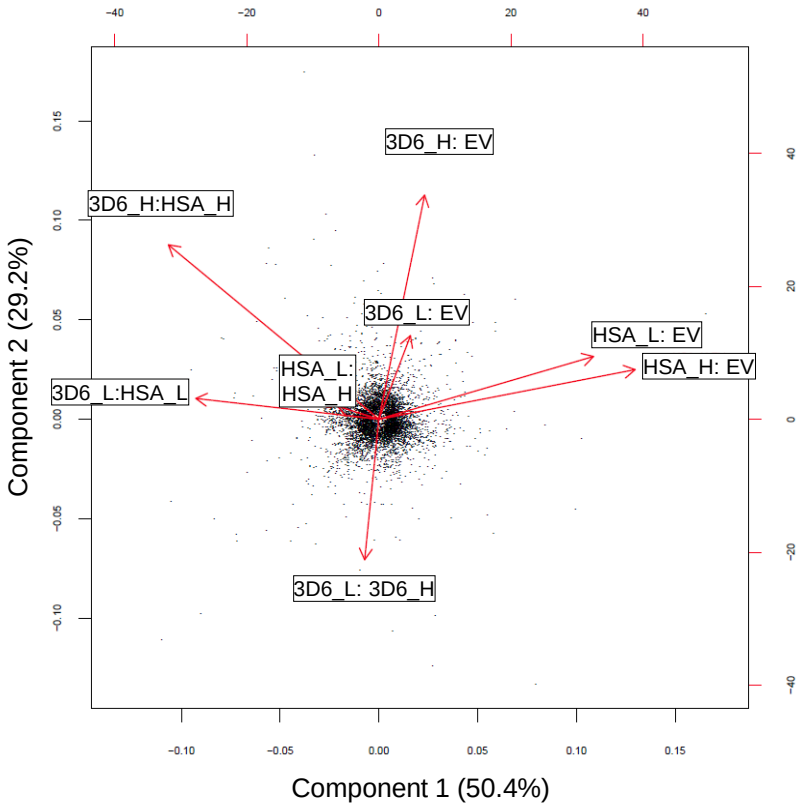


Figure 16. Principal component analysis (PCA) of the whole dataset (8 comparisons). 3D6_H, CHO 3D6scFv-Fc high producer; 3D6_L, CHO 3D6scFv-Fc low producer; HSA_H, CHO HSA high producer; HSA_L, CHO HSA low producer; EV, CHO empty vector (non-producer)

Our results suggest that the reaction of CHO cells to heterologous protein expression is strongly product and/or clone dependent. This would also explain the poor agreement between previously published studies that analyzed the transcriptome of different recombinant CHO cell lines in relation to high q_P (Kang et al. 2014; Seth et al. 2007; Vishwanathan et al. 2014). Different proteins require specific cellular aid with folding, disulfide bond formation or glycosylation depending on their structure. Hence, stress exerted by the high-level expression of heterologous proteins may well initiate different responses in the cell and consequently affect the regulation of gene expression. However, considering the inherent genetic instability of CHO cells, the contribution of clonal variation might not be negligible.

Seth et al. (2007) suggested that hyper-productivity is not the consequence of large expression level changes of a few master genes, but rather the accumulation of many small changes

affecting a wide range of cellular functions including protein secretion, energy metabolism, cell growth/death control and redox balance, and that multiple and alternative routes may lead to high productivity. In a recent study, Kang et al. (2014) analyzed 17 different monoclonal antibody producing cell lines covering a wide range of q_P on transcriptome and proteome level. Using PCA of the global expression data, they found a strong correlation of the primary components to phenotypes such as cell size and proliferation rate but no correlation to productivity, suggesting that only a minor part of the global expression space is related to productivity.

4.2.4 Endoplasmic reticulum (ER) stress and unfolded protein response (UPR)

The endoplasmic reticulum (ER) constitutes a potential bottleneck for the secretion of recombinant proteins. In eukaryotes, the secretion of proteins starts with the translocation of the nascent polypeptide chain into the ER, where several ER-resident proteins including chaperons ensure the correct folding and processing of the protein. Improperly folded proteins are retained in the ER and predominantly subjected to ER-associated degradation (ERAD) which involves the proteasomal degradation after retrotranslocation into the cytoplasm. However, the accumulation of unfolded and misfolded proteins (e.g. due to overexpression of heterologous proteins) results in ER stress and the activation of the unfolded protein response (UPR). ER stress activates the stress sensors ATF6 (activating transcription factor 6), PERK (double-stranded RNA-activated protein kinase-like ER kinase) and IRE1 (inositol requiring enzyme 1), which represent the three branches of the UPR (Figure 17). Each UPR pathway produces a transcription factor ($ATF6_N$, ATF4 and XBP1s, respectively) that induces genes to increase the protein-folding capacity in the ER (chaperons and foldases) and/or enhance the ERAD pathway. In addition, PERK and IRE1 also reduce the protein load entering the ER by attenuating translation or degrading mRNAs, respectively (Walter and Ron 2011). The UPR functions as a feedback loop to mitigate ER stress by reestablishing protein folding homeostasis. Persistent ER stress, however, induces apoptosis. If the protein folding homeostasis cannot be restored, the increased protein synthesis leads to oxidative stress and consequently to cell death (Han et al. 2013).

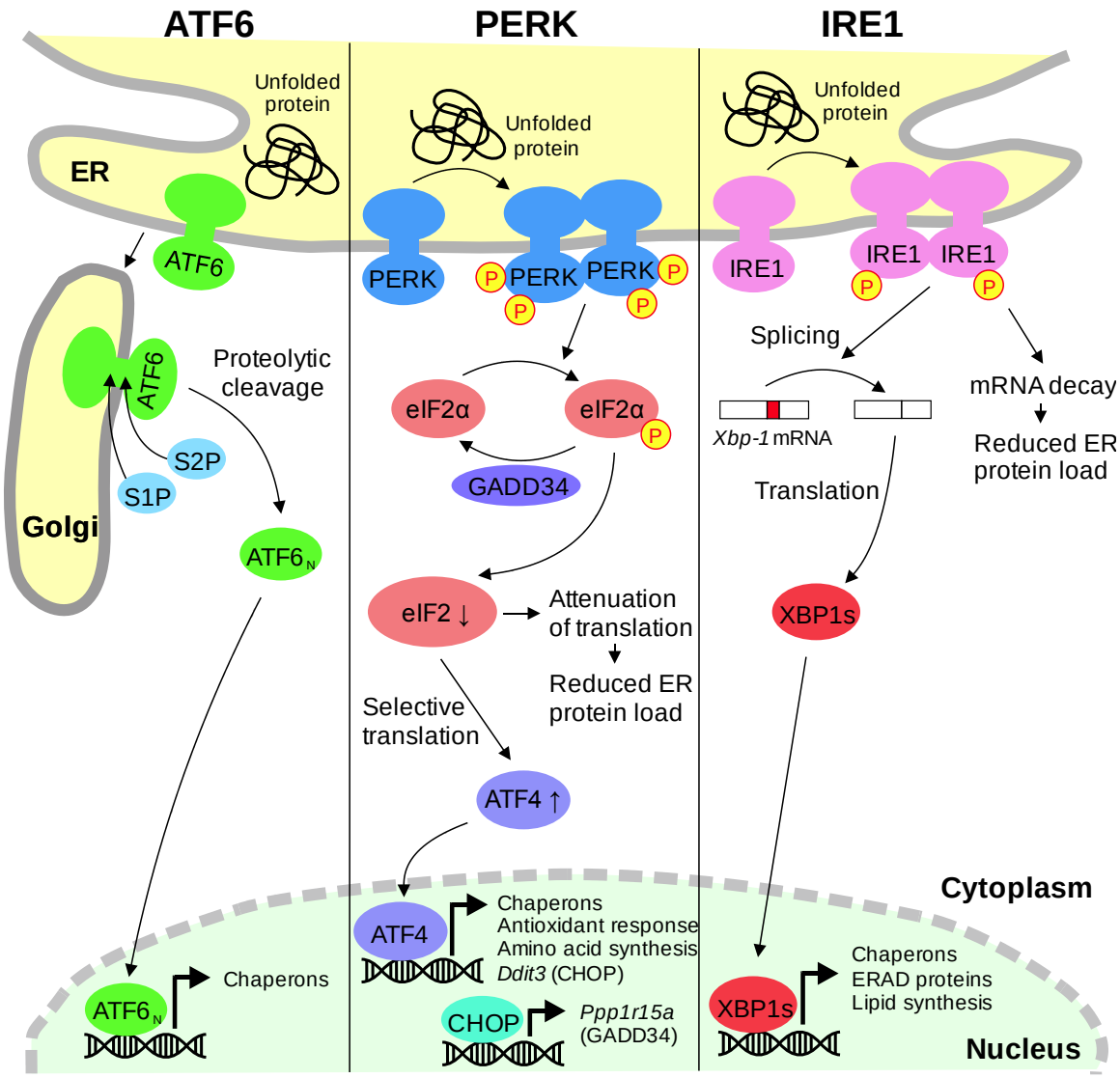


Figure 17. The three branches of unfolded protein response (UPR) in metazoans. ER stress is sensed by three signal transducer (ATF6, PERK and IRE1) resulting in the production of three transcription factors ($ATF6_N$, ATF4 and XBP1s), which drive the transcription of UPR target genes (e.g. chaperones) that mainly serve to increase the protein folding capacity in the ER. PERK and IRE1 also decrease the protein load in the ER by attenuating translation and degrading mRNAs, respectively. Each pathway uses a different mechanism for signal transduction (Proteolysis of ATF6, selective translation of ATF4 and nonconventional splicing of Xbp1).

In order to identify a potential UPR induction, the expression profiles of UPR target genes were analyzed using the microarray data. Potential target genes of $ATF6_N$, ATF4 and XBP1s were retrieved from the literature (Adachi et al. 2008; Han et al. 2013; Okada et al. 2002; Shaffer et al. 2004; Shoulders et al. 2013; Yamamoto et al. 2007).

$ATF6_N$ targets include genes encoding prominent ER-resident proteins involved in protein folding such as BiP/GRP78 (encoded by *Hspa5*), protein disulfide isomerase (PDI, encoded by *P4hb*), HSP90B1/GRP94 (encoded by *Hsp90b1*) and calreticulin (encoded by *Calr*). All

four genes are highly expressed in CHO cells, however, only *Hsp90b1* was slightly up-regulated in the HSA high producer compared to the low producer (Figure 18). *Hspa5* was even down-regulated in all producers compared to the non-producer. In general, an up-regulation of $ATF6_N$ target genes in the producers compared to the non-producer was not observed. *Hmgcs1* (encoding 3-hydroxy-3-methylglutaryl-CoA synthase (HMG-CoA-S)) was up-regulated in both high producers compared to the low producer. However, HMG-CoA-S is part of the cholesterol biosynthesis pathway and the connection to UPR is unclear (Okada et al. 2002). In addition, the genes encoding ATF6 (*Atf6*), S1P (*Mbtps1*) and S2P (*Mbtps2*) were not differentially expressed in any comparison.

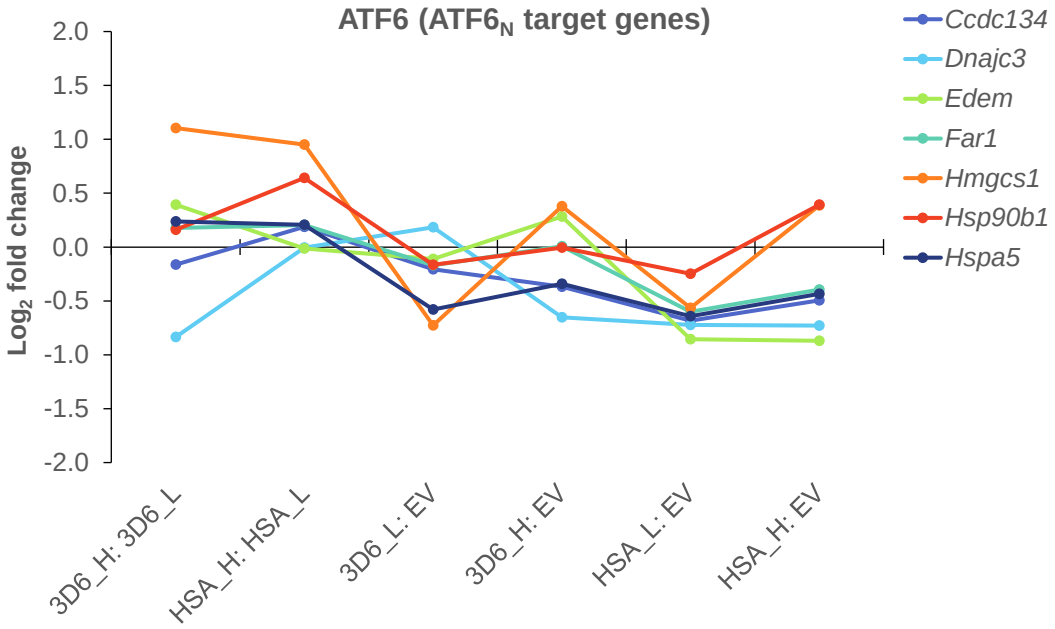


Figure 18. Expression profile of potential $ATF6_N$ target genes (ATF6 branch of UPR). Only significantly differentially expressed genes (FC ≥ 1.5, adj. $p \leq 0.05$) are shown. The potential target genes *Atp2a2*, *Calm1*, *Calr*, *Creld2*, *Derl3*, *Dnajb11*, *Ero1l*, *Ero1lb*, *Galnt3*, *Herpud1*, *Hoxa1*, *Hrg*, *Hyou1*, *Manf*, *Nucb2*, *Ormdl2*, *Os9*, *P4hb*, *Pdia3*, *Pdia4*, *Pdia6*, *Piga*, *Sdf2l1*, *Sel1l*, *Syvn1*, *Tbccd1* and *Wipi1* were not differentially expressed. 3D6_H, CHO 3D6scFv-Fc high producer; 3D6_L, CHO 3D6scFv-Fc low producer; HSA_H, CHO HSA high producer; HSA_L, CHO HSA low producer; EV, CHO empty vector (non-producer)

The expression of the two most important targets of the PERK pathway, *Ddit3* (encoding CHOP) and *Ppp1r15a* (GADD34) was not enhanced because of heterologous protein production. None of the ATF4 target genes was up-regulated in the high producers compared to the low producers (Figure 19).

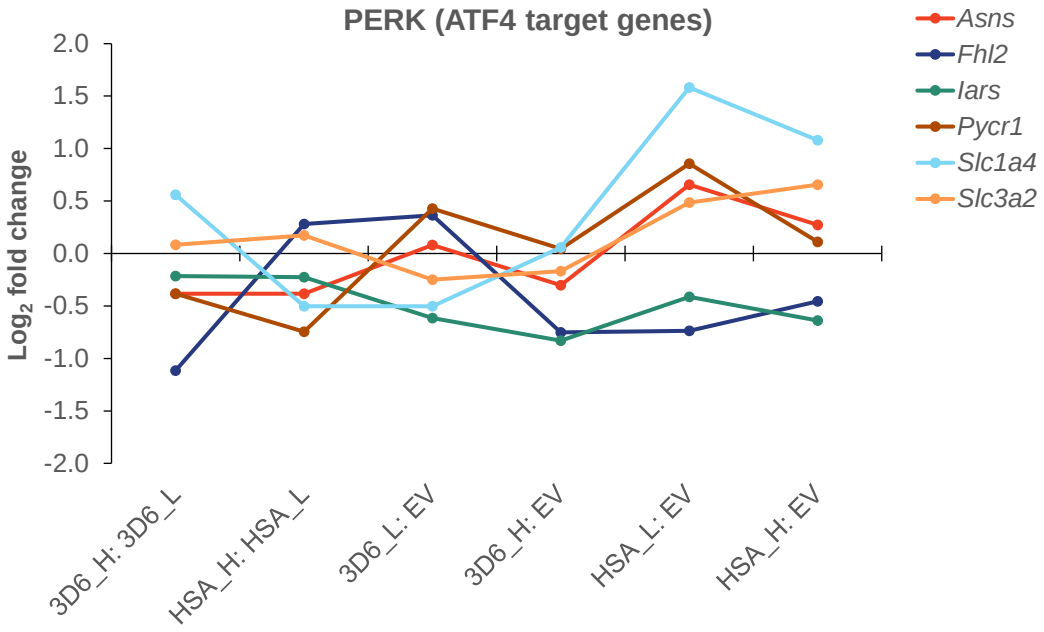


Figure 19. Expression profile of potential ATF4 target genes (PERK branch of UPR). Only significantly differentially expressed genes (FC ≥ 1.5, adj. $p \leq 0.05$) are shown. The potential target genes *Aars*, *Atf3*, *Ddit3*, *Eif2s2*, *Eif3c*, *Eif4g2*, *Eif5*, *Eprs*, *Ero1l*, *Gars*, *Got1*, *Idua*, *Lars*, *Mars*, *Nars*, *Pax6*, *Ppp1r15a*, *Sars*, *Trib3*, *Vars*, *Wars*, *Xpot* and *Yars* were not differentially expressed. 3D6_H, CHO 3D6scFv-Fc high producer; 3D6_L, CHO 3D6scFv-Fc low producer; HSA_H, CHO HSA high producer; HSA_L, CHO HSA low producer; EV, CHO empty vector (non-producer)

Most of the potential target genes of XBP1s were not differentially expressed in any cell line (Figure 20). *Xbp-1* itself was significantly down-regulated in the 3D6scFv-Fc high producer and in both HSA producers compared to the non-producer. *Sec24d* was up-regulated in both high producers compared to the low producer and *Sec31a* was up-regulated in both HSA producers compared to the non-producer. Both genes encode proteins of the COPII complex which is responsible for vesicle budding from the ER and transport to the Golgi apparatus (Jensen and Schekman 2011). Most interestingly, *Serp1* (encoding stress-associated endoplasmic reticulum protein 1 (SERP1)) was up-regulated in the 3D6scFv-Fc high producer compared to the low producer as well as the non-producer. SERP1 is identical to the ribosome-associate membrane protein 4 (RAMP4), a membrane protein associated with the Sec61 complex (ER translocon), and the deletion of SERP1/RAMP4 was shown to induce ER stress (Hori et al. 2006). SERP1/RAMP4 also controls glycosylation of the major histocompatibility complex class II-associated invariant chain (Schröder et al. 1999). Yamaguchi et al. (1999) demonstrated that SERP1/RAMP stabilizes newly synthetized membrane proteins during ER stress and subsequently facilitates glycosylation. Hence,

SERP1 might be involved in processing and glycosylation of the 3D6scFv-Fc antibody in the ER.

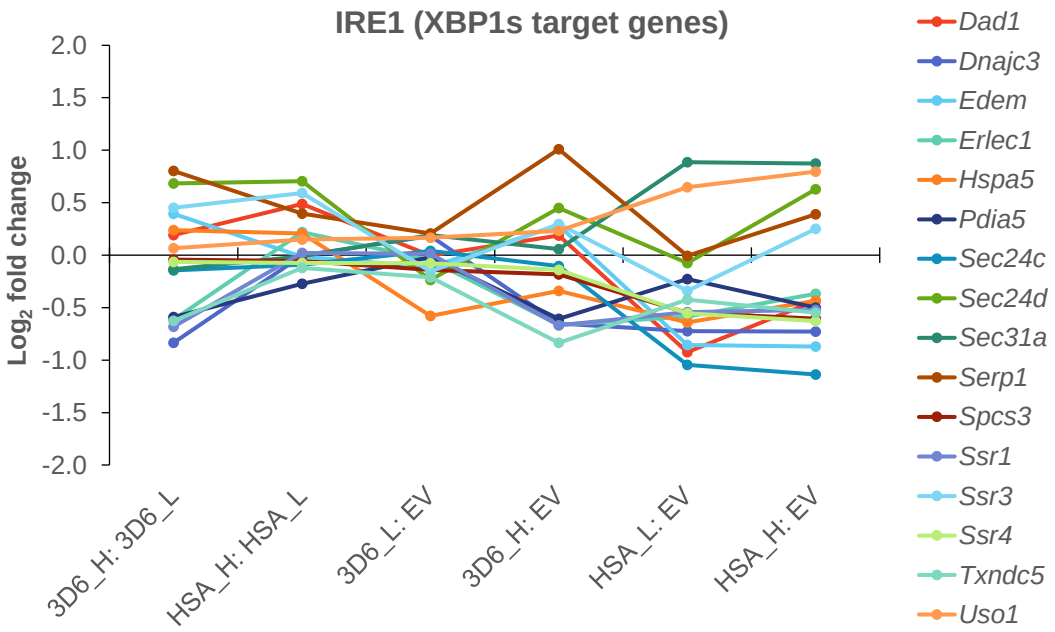


Figure 20. Expression profile of potential XBP1s target genes (IRE1 branch of UPR). Only significantly differentially expressed genes (FC ≥ 1.5, adj. $p \leq 0.05$) are shown. The potential target genes *Arcn1*, *B4galt2*, *Calr*, *Canx*, *Cope*, *Creld2*, *Ddost*, *Dnajb11*, *Dnajb9*, *Edem2*, *Edem3*, *Ero1lb*, *Fkbp11*, *Fut8*, *Golgb1*, *H13*, *Herpud1*, *Hyou1*, *Lman1*, *Man1b1*, *Mgat2*, *Mogs*, *Os9*, *P4hb*, *Pdia2*, *Pdia3*, *Pdia4*, *Pdia6*, *Ppib*, *Rhoq*, *Rpn1*, *Sec23b*, *Sec61a1*, *Sec61g*, *Sel1l*, *Srp54b*, *Srpr*, *St6galnac4*, *Stt3a*, *Sulf1*, *Syvn1*, *Tram1* and *Wipi1*were not differentially expressed. 3D6_H, CHO 3D6scFv-Fc high producer; 3D6_L, CHO 3D6scFv-Fc low producer; HSA_H, CHO HSA high producer; HSA_L, CHO HSA low producer; EV, CHO empty vector (non-producer).

Nevertheless, a clear induction of UPR was not observed in any of the producing cell lines. Hence, the high-level expression of 3D6scFv-Fc and HSA did not induce ER stress in the cell lines analyzed in the present study, which is also supported by the qRT-PCR analysis of the product mRNA levels (Figure 21). For both model proteins, an ~8-fold higher product transcript level in the high producer compared to the low producer resulted in an 8-fold higher q_P, suggesting that secretion does not constitute a bottleneck in these cell lines. Consequently, a cell engineering strategy aiming to enhance the secretory capacity would most likely not increase recombinant protein production. This might also explain the different outcomes of cell engineering studies targeting chaperons and the UPR signaling pathway (see Chapter 1.4), suggesting that the effectiveness of these approaches depends on the particular recombinant protein and its expression level.

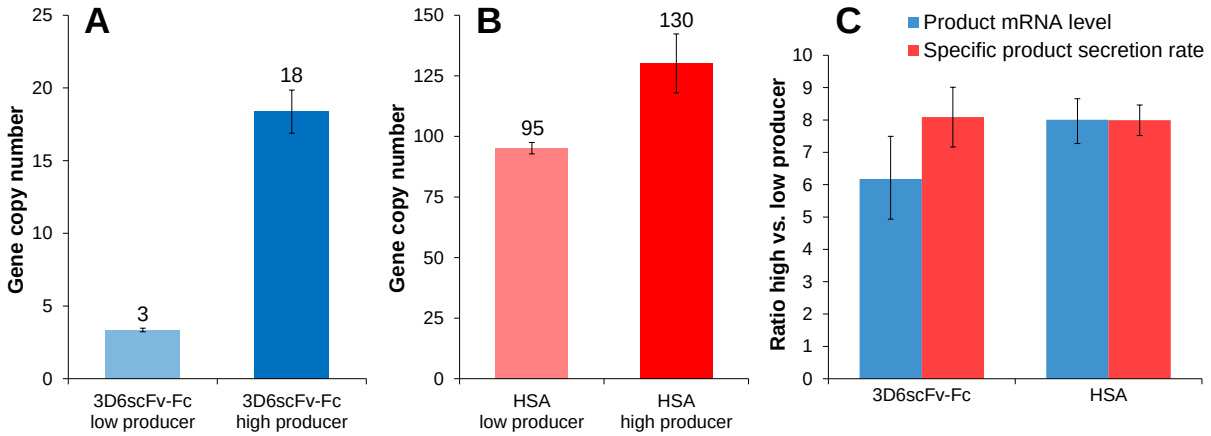


Figure 21. Correlation of transgene copy number, product mRNA level and specific product secretion rate (q_P). Transgene copy number of (**A**) CHO 3D6scFv-Fc low and high producer, and (**B**) HSA low and high producer. Gene copy numbers were determined by qPCR specific for the cytomegalovirus (CMV) major immediate-early (MIE) enhancer/promoter (absolute values per unreplicated genome). Data represent mean values ± SD (n = 2). (**C**) Ratios of product mRNA and q_P of 3D6scFv-Fc and HSA high producer to low producers. Transcript levels were determined by qRT-PCR. Data represent mean values ± SE of three independent steady-state cultivations.

4.2.5 Engineering transcription efficiency

The perfect correlation of transgene copy number, product mRNA level and q_P (Figure 21) indicates that there is most likely no limitations at the stage of translation, protein processing or secretion in the recombinant CHO cell lines investigated in this study. Hence, the efficiency of transgene transcription driven by the human cytomegalovirus (CMV) major immediate-early (MIE) enhancer/promoter may potentially constitute a bottleneck in these cell lines.

The CMV MIE enhancer/promoter contains binding sites for various transcription factors including NF-κB, CREB/ATF, SP1, YY1, AP-1 and NFI (Stinski and Isomura 2008). Some of them were also differentially expressed in our study (Figure 22).

The nuclear factor-kappa B (NF-κB) family is composed of two subfamilies (NF-κB and Rel). Two members of the NF-κB subfamily (NF-κB1 and NF-κB2) are synthetized as large precursor (p105 and p100) and processed to become the active DNA-binding proteins p50 and p52, respectively (Gilmore 2006). RelA, RelB and c-Rel are the most prominent members of the Rel subfamily. Various studied have demonstrated that NF-κB is a key regulator of transcription from the human CMV MIE enhancer/promoter, although this remains controversial as it has also been shown that NF-κB binding sites are not required for gene expression (Gustems et al. 2006). We found *Nfkb2* (encodes NF-κB2) exclusively up-

regulated in the 3D6scFv-Fc high producer, and *Relb* (encodes RelB) exclusively up-regulated in the HSA high producer (Figure 22).

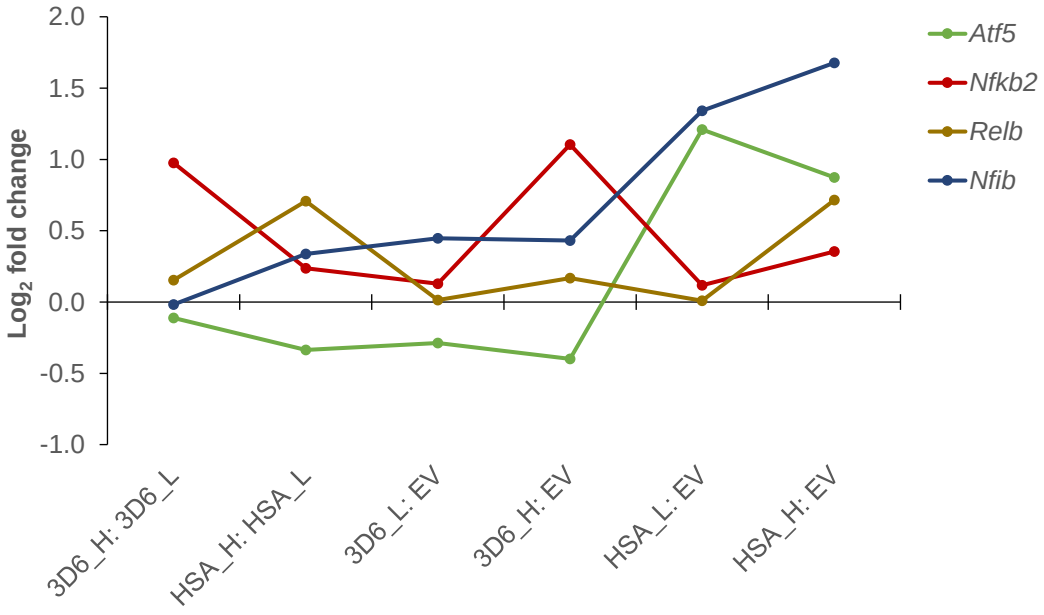


Figure 22. Expression profile of transcription factors associated with the cytomegalovirus (CMV) major immediate-early (MIE) enhancer/promoter. Only significantly differentially expressed genes (FC ≥ 1.5, adj. $p \leq 0.05$) are shown. *Atf2*, *Atf3*, *Atf4*, *Atf5*, *Creb1*, *Jun*, *Nfia*, *Nfix*, *Rel*, *Rela*, *Sp1* and *Yy1* were not differentially expressed. *Atf5* encodes Activating transcription factor 5 (ATF5). *Nfkb2* encodes nuclear factor-kappaB p100 subunit (NF-κB2). *Relb* encodes RelB a transcription factor of the NF-κB family. *Nfib* encodes nuclear factor I/B (NFIB). 3D6_H, CHO 3D6scFv-Fc high producer; 3D6_L, CHO 3D6scFv-Fc low producer; HSA_H, CHO HSA high producer; HSA_L, CHO HSA low producer; EV, CHO empty vector (non-producer)

Nfib, which encodes nuclear factor I/B (NFIB), was up-regulated in the HSA low producer (2.5-fold) and HSA high producer (3.2-fold) compared to the non-producer, whereas *Nfib* was not significantly differentially expressed in both 3D6scFv-Fc producers (Figure 22). The HSA producing clones had considerably higher absolute transgene (CMV promoter) copy numbers than the 3D6scFv-Fc producing clones (Figure 21A and B). Thus, *Nfib* up-regulation might be the consequence of the considerably higher number of NFIB binding sites. In a recent study, it was observed that the overexpression of NFIB led to a strong up-regulation of gene expression driven by a CMV promoter in HEK293 cells (Liang et al. 2012). Hence, it appears that NFIB is a crucial transcription factor of the CMV promoter and therefore a promising target for cell engineering to enhance specific productivity of recombinant CHO cells.

Using hierarchical cluster analysis, we identified *Nfkbie* and *Mdfic* showing a similar expression profile as *Nfib*. *Nfkbie* encodes IκBε which is an inhibitor of different NF-κB

proteins including RelA, c-Rel, p50 and p52 (Li and Nabel 1997). The human *Mdfic* gene which encodes the human I-mfa domain containing protein (HIC), was shown to activate transcription elongation (Young et al. 2007). Interestingly, not the protein but the 3' UTR of HIC mRNA binds and activates the positive transcription elongation factor b (P-TEFb), which phosphorylates the carboxy-terminal domain (CTD) of RNA polymerase II in order to stimulates transcription. In CHO cell lines producing human Epo and IFNγ, it was shown that the overexpression of the 3' UTR of HIC mRNA, but not the full-length HIC mRNA, significantly improves recombinant protein production (Liu et al. 2009).

In conclusion, targeting key components that control the transcription efficiency of the gene of interest might be a promising strategy to increase specific productivity of cell lines where translation, protein processing or secretion are not limiting.

4.3 Global miRNA expression profiling in recombinant CHO cell lines

In this study, the microRNA (miRNA) expression profiles of high, low and non-producing recombinant CHO cell lines were analyzed and compared in order to identify miRNAs that are involved in heterologous protein synthesis and secretion. Therefore, CHO cell lines expressing 3D6scFv-Fc and HSA were generated and low and high producers were selected at different stages of transgene amplification (Publication B and C). In addition a mock transfected CHO cell lines was established (non-producer). All five cell lines were cultivated in biological triplicates in 800 mL DASGIP bioreactors under steady-state conditions using a chemostat process. Cross-species miRNA microarrays (human, mouse and rat) were used as screening tools and quantitative reverse transcription polymerase chain reaction (qRT-PCR) for the validation of differentially expressed miRNAs (Publication C). In the following, the most important findings of Publication C as well as additional aspects of this study are summarized and discussed.

4.3.1 Product/clone dependency of miRNA expression

In order to identify differences in the miRNA expression profiles of the five CHO cell lines analyzed in this study, following eight comparisons were made.

- CHO 3D6scFv-Fc high producer vs. CHO 3D6scFv-Fc low producer (3D6_H: 3D6_L)
- CHO HSA high producer vs. CHO HSA low producer (HSA_H: HSA_L)
- CHO 3D6scFv-Fc low producer vs. CHO HSA low producer (3D6_L: HSA_L)
- CHO 3D6scFv-Fc high producer vs. CHO HSA high producer (3D6_H: HSA_H)
- CHO 3D6scFv-Fc low producer vs. CHO empty vector (3D6_L: EV)
- CHO 3D6scFv-Fc high producer vs. CHO empty vector (3D6_H: EV)
- CHO HSA low producer vs. CHO empty vector (HSA_L: EV)
- CHO HSA high producer vs. CHO empty vector (HSA_H: EV)

MicroRNAs with a fold change (FC) ≥ 1.5 and an adjusted *p*-value (adj. *p*) ≤ 0.05 (multiple testing correction based on Benjamini and Hochberg) were considered as differentially expressed. In general, higher miRNA expression levels were observed in the cell lines expressing heterologous proteins compared to the non-producer, suggesting that miRNAs play an important role in regulation of processes related to recombinant protein synthesis and secretion. Several differentially expressed miRNAs were identified between the producers and

the non-producer as well as between 3D6scFv-Fc producer and HSA producer, but only few between the high and low producers (Table 2).

Table 2. Number of differentially expressed miRNAs (Fold change ≥ 1.5 and adj. $p \leq 0.05$)

	3D6_H: 3D6_L	**HSA_H: HSA_L**	**3D6_L: HSA_L**	**3D6_H: HSA_H**	**3D6_L: EV**	**3D6_H: EV**	**HSA_L: EV**	**HSA_H: EV**
Up	2	1	26	20	30	18	7	13
Down	7	6	4	9	0	6	12	18
Total	9	7	30	29	30	24	19	31

However, none of the differentially expressed miRNAs was commonly up- or down-regulated comparing high to low producer for both model proteins (Figure 23A). To increase the statistical power (reduce the false negative rate), the data was reanalyzed without applying multiple testing correction, confirming that there are no miRNAs commonly differentially expressed (FC > 1.5 and $p < 0.05$) between high and low producer for HSA and 3D6sc-Fv-Fc. Comparing 3D6scFv-Fc low and high producer against the non-producer showed that only seven out of 30 differentially expressed miRNAs in the low producer were also up- or down-regulated in the high producer (Figure 23B). For HSA, the overlap was considerably higher with 14 out of 19 (Figure 23C). Interestingly, all commonly differentially expressed miRNAs in the 3D6scFv-Fc producers were up-regulated (Figure 23B), whereas for HSA most of them (10 out of 14) were down-regulated (Figure 23C). The Venn diagram of both low producers as well as both high producers versus the non-producer showed only a very little overlap in both cases (Figure 23D and E). Comparing all producers against the non-producer showed that only one miRNA was commonly differentially expressed (Figure 23F).

The results of the miRNA expression profiling experiment are consistent with those of the mRNA profiling experiment (see Chapter 4.2.3), supporting the hypothesis that the reaction of CHO cells to heterologous protein expression is highly product and/or clone dependent. Consequently, cell engineering approaches to improve the productivity of recombinant CHO cell lines may as well be product and/or clone-specific and not generally applicable.

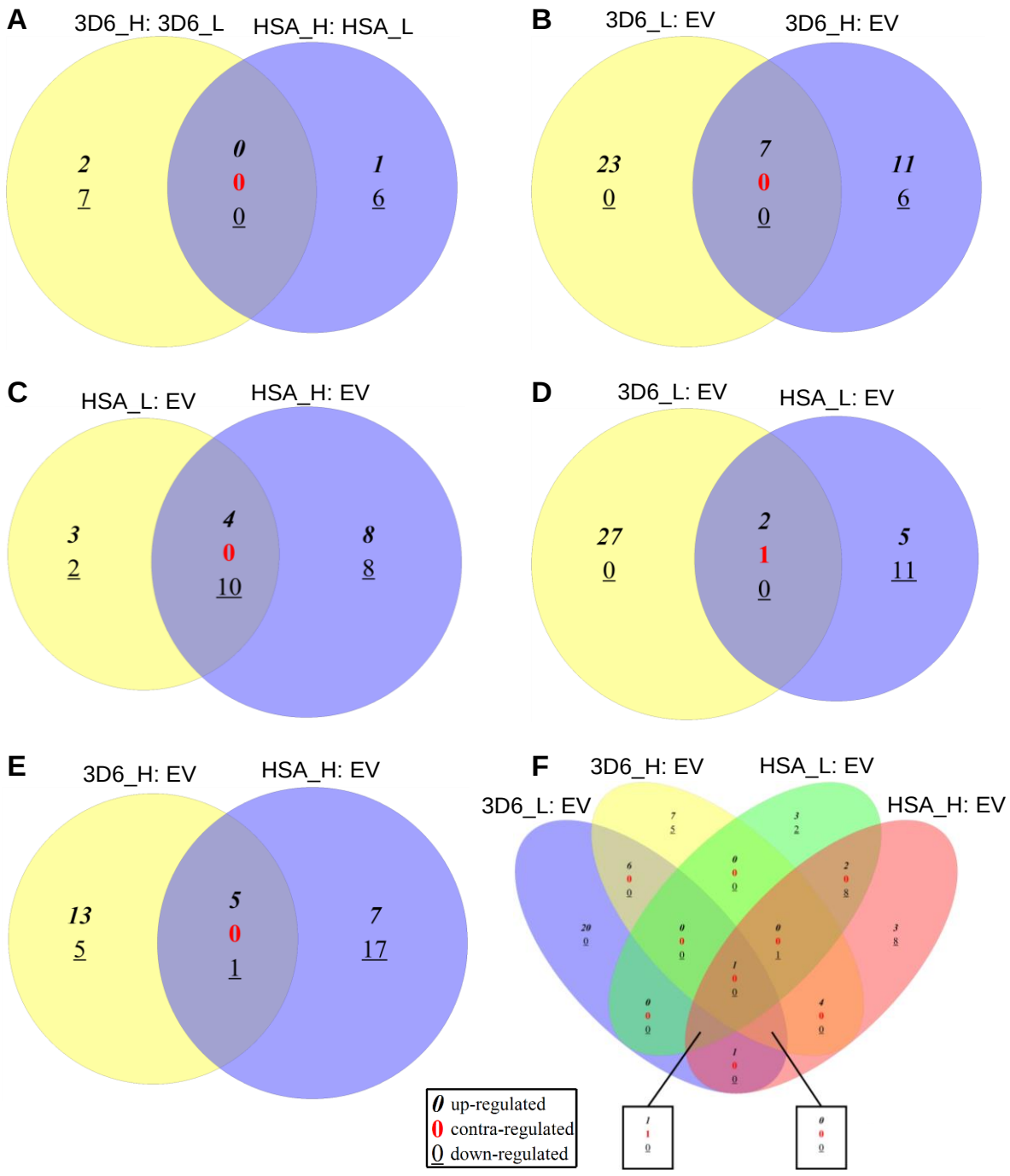


Figure 23. Number of commonly and exclusively differentially expressed miRNAs. 3D6_H, CHO 3D6scFv-Fc high producer; 3D6_L, CHO 3D6scFv-Fc low producer; HSA_H, CHO HSA high producer; HSA_L, CHO HSA low producer; EV, CHO empty vector (non-producer). Venn diagrams were generated using VennPlex (Cai et al. 2013).

4.3.2 Impact of confirmed miRNAs

The expression profiles of 11 priority miRNAs were confirmed by qRT-PCR. Some of them were already found to be differentially expressed between producer and non-producer in similar previously published studies. Up-regulation of miR-10b-5p and down-regulation of miR-21-5p was observed in an Epo-Fc and an antibody producer compared to the respective parental CHO cell line (Hackl et al. 2011), which is in compliance with the results of the HSA producers but not the 3D6scFv-Fc producers. In another study, Lin et al. (2011) profiled miRNA expression in four recombinant CHO cell lines expressing the same human IgG and compared them with the parental DG44 cell line. In accordance with our results, they found miR-221-3p being significantly down-regulated in the recombinant cell lines. They also observed a down-regulation of miR-125b-5p in two clones, as observed in the 3D6scFv-Fc high producer in our study. However, they detected opposed effects for miR-19a-3p and let-7b-5p.

It was already shown that the overexpression of miR-21-5p reduces the specific productivity of recombinant CHO cells (Jadhav et al. 2012), consequently a knockdown of miR-21-5p might increase the specific productivity. As one of the most abundant and best studied miRNAs, miR-21-5p has been linked to cell proliferation, apoptosis, migration (Krichevsky and Gabriely 2009) and cellular longevity (Dellago et al. 2013). *Nfib* and *Pdcd4* are well described targets that were also identified as potential targets of miR-21-5p in CHO cells in our study. Programmed cell death protein 4 (encoded by *Pdcd4*) acts as tumor-suppressor which is up-regulated during apoptosis (Shibahara et al. 1995), and it was shown to inhibit translation initiation by binding to eIF4A (Suzuki et al. 2008). However, this cannot explain the observed alteration in specific productivity. The transcription factor Nuclear factor I/B (encoded by *Nfib*) is a transcriptional repressor of miR-21 which is involved in a double negative feedback mechanism to sustain the level of miR-21 transcription (Fujita et al. 2008). Thus, the down-regulation of miR-21-5p in the HSA producers may only be a consequence of the increased *Nfib* levels, which we hypothesized to be caused by the high transgene copy number of these cell lines (see Chapter 4.2.5).

Like miR-21-5p, miR-10b-5p was described as an oncogenic miRNA with a pro-proliferative and anti-apoptotic function (Lin et al. 2012). We identified *Sdc1* as a potential target of miR-10b-5p, which was also described in breast cancer cells (Ibrahim et al. 2012). Its gene product Syndecan-1 is a transmembrane proteoglycan which acts as a co-receptor for various

growth factors and growth factor receptors, thereby having a crucial effect on processes involved in the regulation of cell growth (Szatmari and Dobra 2013).

In human hepatocellular carcinoma cells, miR-221-3p was found to control the expression of cyclin-dependent kinase inhibitor 1B ($p27^{Kip1}$) and cyclin-dependent kinase inhibitor 1C ($p57^{Kip2}$), which are important cell cycle inhibitors (Fornari et al. 2008). *Cdkn1b* which encodes $p27^{Kip1}$ was also identified as a potential target of miR-221-3p in our study. Increased specific productivity in CHO cells was achieved by conditional overexpression of $p27^{Kip1}$, which induced a G1-specific growth arrest (Meents et al. 2002b). Hence, knockdown of miR-221 might have a similar effect, however, transient overexpression of miR-221 had no significant effect on growth and productivity in an Epo-Fc producing CHO cell line (Jadhav et al. 2012). In human hepatocellular carcinoma cells, miR-221 was also linked to apoptosis regulation, where ER stress-induced apoptosis was attenuated by miR-221/222 inhibitors (Dai et al. 2010).

Diverse functions have been described for miR-125b, as it can act as tumor suppressor or as oncogene and it was shown to promote apoptosis in various human cell lines by suppressing the expression of Bcl-2 family members (Gong et al. 2013). Hence, the strong down-regulation of miR-125b-5p in the 3D6scFv-Fc high producer might be a response to cellular stress caused by high expression of the scFv-Fc antibody in order to prevent apoptosis. However, in human keratinocytes it was also demonstrated that miR-125b modulates cell proliferation by targeting *Fgfr2* which encodes fibroblast growth factor receptor 2 (Xu et al. 2011a). *Fgfr2* was also identified as potential target in CHO cells in our study.

Two highly expressed members of the let-7 family, let-7b-5p and let-7c-5p, were up-regulated in all producers relative to the non-producer and their expression levels correlated with specific productivity. The let-7 family is highly conserved across animals and has been associated with the regulation of various cellular processes including cell differentiation and proliferation as well as tumor suppression (Roush and Slack 2008). In human cancer cells, the overexpression of let-7b led to reduced Dicer protein levels (but not mRNA levels) as well as significant inhibition of cell proliferation (Jakymiw et al. 2010). Dicer expression was also found to be strongly correlated with growth rate in CHO cells (Hackl et al. 2014). Hence, let-7b might as well be able to modulate cell proliferation in CHO cells. In mice, it was found that let-7b also regulates glucose metabolism by targeting genes of the PI3K/AKT/mTOR pathway (Zhu et al. 2011). We identified *Igf2bp* as a potential target of let-7b-5p in CHO cells. Insulin-like growth factor 2 mRNA binding protein 3 (encoded by *Igf2bp*) is a

translational activator of insulin-like growth factor 2 (encoded by *Ifg2*) which is an activator of the MAPK as well as the PI3K/AKT/mTOR pathway (Suvasini et al. 2011).

In various cancer studies, miR-19a was described as an oncogenic miRNA with a pro-proliferative and anti-apoptotic function, and in laryngeal squamous cell carcinoma, it was shown to target *Timp2* (Wu et al. 2014), which was also identified as potential target of miR-19a-3p in our study. Tissue inhibitor of metalloproteinase 2 (TIMP-2), which is encoded by *Timp2*, plays an important role in the suppression of cell proliferation. In human endothelial cells, shear stress induces miR-19a that targets the cyclin D1 mRNA, which leads to an arrest of cell cycle at G1/S transition (Qin et al. 2010). In a recent study, miR-19a was also linked to the PI3K/AKT/mTOR pathway, suggesting a positive feedback loop (He et al. 2013).

The two members of the miR-99 family, miR-99a and miR-100, have been described to function as tumor suppressor or oncogene. In various cancer studies, it was demonstrated that overexpression of miR-99a/100 inhibits cell proliferation and/or promotes apoptosis by suppressing the expression of mTOR (Chen et al. 2012; Cui et al. 2012; Sun et al. 2013). However, in myeloid leukemia, it was shown that miR-99a/100 play an oncogenic role by regulating the expression of CTD small phosphatase-like protein (CTDSPL) and tribbles pseudokinase 2 (TRIB2), which are two tumor suppressors (Zhang et al. 2013; Zheng et al. 2012). *Ctdspl* was also identified as potential target of miR-99a-5p in CHO cells in our study.

Diverse functions have also been described for miR-193a. In leukemia cells, miR-193a acted as a tumor suppressor by inducing G1 growth arrest and apoptosis (Li et al. 2013). However, in mesenchymal stem cells (MSCs), it was shown that miR-193a promotes proliferation by targeting the mRNA of inhibitor of growth family 5 (ING5) which inhibits cyclin-dependent kinase 2 (CDK2), but does not affect apoptosis (Wang et al. 2012).

In rat cardiomyocytes, it was shown that the overexpression of miR-350 induces hypertrophy and apoptosis by repression the c-Jun N-terminal kinase (JNK) and the p38 mitogen activated protein kinase (MAPK) pathway (Ge et al. 2013).

All miRNAs described above, have been related to diverse functions mainly associated with cell proliferation and/or apoptosis, so that their relation to productivity may only be indirect and therefore difficult to interpret. Interestingly, let-7b, miR-19a and the miR-99 family have been associated with the PI3K/AKT/mTOR pathway. It has already been showed that the ectopic expression of human mTOR can improve cell growth, proliferation, viability, robustness and specific productivity of recombinant CHO cells (Dreesen and Fussenegger 2011), consequently modulating expression of let-7b, miR-19a or the miR-99 family may also

affect bioprocess-relevant phenotypes. Nevertheless, it is necessary to further evaluate the identified target miRNAs experimentally. This can be done by transient and stable overexpression of the miRNA as well as knockdown using either antisense miRNA oligonucleotides (antagomirs) or miRNA sponges (Jadhav et al. 2013).

4.3.3 Functional characterization of miRNAs

Identifying targets mRNAs is crucial to understand the biological function of miRNAs. Biological and biochemical methods for high-throughput miRNA target identification are very labor-intensive and time-consuming, hence, reliable software prediction tools would be very beneficial. However, today's computational target prediction tools are inefficient due to the complexity and diversity of miRNA-mRNA interactions. In addition, no prediction tools as well as experimentally validated target databases that are specific for CHO cells are currently available.

Various web-based tools such as DIANA miRPath (Vlachos et al. 2012), miRSystem (Lu et al. 2012) and miRGator (Cho et al. 2013) have been developed for the functional characterization of miRNAs based on gene enrichment analysis. They heavily rely on correct computational miRNA target prediction or comprehensive validated target databases. But experimentally validated target databases are currently incomplete, with the best coverage for human, but a low number of validated targets in other organism. For instance, miRTarBase 4.5 (Hsu et al. 2014) contains 489 validated human targets of the very well-studied miR-21-5p, whereas only 11 targets are present for mouse and nine for rat. Additionally, most validated targets originate from cancer research, consequently interactions relevant for cell culture technology might be underrepresented. Considering also the significant false-positive and false-negative rates of target prediction algorithms, it is very likely that results of the currently available functional analysis tools are biased and therefore might lead to misinterpretations.

5. Conclusion and perspectives

Chinese hamster ovary (CHO) cells are currently the workhorses of the biopharmaceutical industry for the production of recombinant therapeutic proteins. However, due to recent advances in glycoengineering, yeasts such as *Pichia pastoris* are about to enter this field. Although, a large amount of information about heterologous protein expression has been published for both expression systems, a quantitative comparison of these data is very difficult, as they were generated in non-comparable ways following different research questions. In addition, the productivity of the individual expression systems strongly depend on the particular protein, however, proteins produced in yeasts are generally different ones than those expressed in mammalian cells. In order to conduct a thorough system comparison between CHO cells and *P. pastoris*, we established comparable high producer cell lines and strains that express two structurally different model proteins. One of them, human serum albumin (HSA), can be produced at very high levels in *P. pastoris*, whereas the other one, a single chain Fv-Fc fusion antibody (3D6scFv-Fc), is more complex and contains the Fc-specific N-glycosylation site. In *P. pastoris*, the biomass-specific secretion rate (q_P) was significantly lower (~40 times) for 3D6scFv-Fc than for HSA, whereas q_P was similar for both model proteins in CHO cells. Comparing both organisms, the q_P of the CHO cell lines was more than 1000-fold higher for 3D6scFv-Fc and 26-fold higher for HSA, which clearly demonstrates that the protein secretion machinery is much more efficient in CHO cells than in *P. pastoris*. The cellular mechanisms and processes that control and limit recombinant protein synthesis and secretion in mammalian cells as well as in yeast are still poorly understood. Hence, a comparative genome-wide analysis of the established CHO cell lines and *P. pastoris* strains using various omics approaches may provide valuable insights, which enable the identification of conserved and species-unique mechanisms and pathways that are involved in recombinant protein production. The analysis of differentially regulated conserved pathways might than be useful to identify bottlenecks of heterologous protein secretion in *P. pastoris*.

In order to investigate recombinant protein production in CHO cells and to provide transcriptomics and proteomics data for a cross-species comparison with *P. pastoris*, we established low and high producing CHO cell lines expressing HSA and the 3D6scFv-Fc antibody as well as a non-producer. A prerequisite for a comparative physiological analysis of different cell lines is the generation of comparable and defined conditions in order to obtain reproducible and meaningful results. In batch or fed-batch cultures the physico-chemical

conditions are very dynamic and have a significant impact on the cell's transcriptome (Hernandez Bort et al. 2012). For this reason, we applied steady-state cultivation using a chemostat. As the standard industrial bioprocess is fed-batch, chemostat cultures have not been applied in any gene, miRNA or protein expression profiling study of CHO cells yet. Accordingly, this is the first report of transcriptomics data which were generated under steady-state conditions.

Due to the lack of publicly available CHO sequence information, our previous microarray analyses of CHO cells have relied on the cross-species hybridization to commercially available microarrays designed for closely related species such as mouse and rat (Ernst et al. 2006). For the present study, we designed a novel CHO-specific microarray based on the first published genomic sequence of a CHO cell line (Xu et al. 2011b). A comprehensive validation of the obtained microarray results using qRT-PCR confirmed that our CHO-specific microarray provided accurate and reliable data.

A multitude of genes were differentially expressed between high, low and non-producing cell lines. However, we observed a considerable divergence in differentially regulated genes that are potentially related with recombinant protein biosynthesis and secretion between cell lines producing different proteins. This might be caused by the different structure and properties of HSA and the 3D6scFv-Fc antibody, however, the contribution of effects caused by clonal variation is unknown and may not be negligible. CHO cells are characterized by a high degree of genetic and phenotypic diversity and so each cell line has its own unique genome, transcriptome and proteome (Wurm 2013). Hence, the comparison of expression data of different cell lines may be very difficult. The elimination of effects that are caused by clonal variation would require the analysis of several independent clones that show similar bioprocess-relevant phenotypes (growth rate, specific productivity and longevity). However, this would be a very time-consuming, labor- and cost-intensive study, especially when conducted for several different proteins including high and low producers.

The gene expression profiling data of this study were also systematically analyzed using hierarchical and *k*-means clustering as well as Gene Ontology (GO) and KEGG pathway enrichment analysis. But the identified enriched biological processes, molecular functions and pathways could not directly be related to protein synthesis or secretion (data not shown). However, our results support the hypothesis of Seth et al. (2007), which suggests that high productivity may not be ascribed to large changes of a few master genes, but rather to minute

changes affecting a multiplicity of cellular functions including energy metabolism, redox balance, protein processing and secretion as well as growth and death control.

Although we could not increase the productivity of the high producing cell lines by increasing the selection pressure (MTX concentration), our results strongly indicate that there are no limitations at the stage of translation, protein processing or secretion, as the product mRNA levels perfectly corresponded to the increase of q_P between high and low producers. In addition, we did not observe an induction of the unfolded protein response (UPR) and especially the genes encoding the prominent chaperons BiP/GRP78, PDI, calreticulin and calnexin were not affected or even down-regulated upon recombinant protein overexpression. Hence, in the cell lines of this study, the limitations of recombinant protein production are rather at the stage of transcription. Various transcription factors have been reported that modulate transcription efficiency of the cytomegalovirus (CMV) major immediate-early (MIE) enhancer/promoter, however, they have not been used as targets for cell engineering of CHO cells yet. Therefore, we suggest that the overexpression of transcription factors may be a promising strategy to enhance recombinant protein production of CHO cell lines experiencing no protein processing or secretion bottleneck.

In addition, we analyzed the microRNA (miRNA) expression profiles using a cross-species miRNA microarray as well as qRT-PCR. MicroRNAs are short (~22 nt) endogenous RNAs that play an important role in the regulation of gene expression (Bartel 2004), and they have been considered as promising targets for pathway engineering of CHO cell factories, because they regulate entire gene networks and do not constitute an additional burden to the cell's translational machinery (Hackl et al. 2012b; Müller et al. 2008).

Our results indicate that miRNAs play an important role in recombinant cell lines, as producing cell lines appear to have higher levels of mature miRNAs in general. However, the comparison of miRNA expression profiles of different cell lines, both from this study and previously published results, revealed little consensus. This supports the findings of the mRNA profiling study, indicating that the reaction of CHO cells to overexpression of heterologous proteins is strongly protein and/or clone dependent.

The confirmed priority miRNAs have been related to diverse functions mainly associated with cell proliferation and/or apoptosis, so that their relation to productivity may only be indirect and therefore difficult to interpret. Interestingly, let-7b, miR-19a and the miR-99 family have been associated with the PI3K/AKT/mTOR signaling pathway in various cancer studies. It has already been showed that the overexpression of mTOR can improve cell growth,

proliferation, viability, robustness and specific productivity of recombinant CHO cells (Dreesen and Fussenegger 2011), consequently modulating expression of let-7b, miR-19a or the miR-99 family may also affect bioprocess-relevant phenotypes.

The identification of miRNA-mRNA interactions is essential to understand the biological function of miRNAs. Experimental screening methods are very labor-intensive and time-consuming, however, methods based on computational prediction and correlation of miRNA and mRNA expression data are very inefficient due to the complexity of miRNA regulatory mechanisms. Statistical analysis of the identified miRNA-mRNA interactions indicated a considerable rate of false positives. Our results highlight the urgent need for reliable CHO-specific miRNA target prediction tools and experimentally validated target databases in order to facilitate functional analysis of high-throughput miRNA expression data in CHO cells.

As a next step, protein expression results which were generated using MS-based shotgun proteomics need to be included in order to get a more comprehensive and clearer picture about the affected processes. In addition, the protein expression data will be very helpful to improve miRNA target identification.

Furthermore, overexpression and knockdown experiments should be done in order to investigate the biological function of the identified genes and miRNAs in CHO cells as well as their impact on bioprocess-relevant properties such as cell proliferation, protein synthesis/secretion and longevity.

6. References

Adachi Y, Yamamoto K, Okada T, Yoshida H, Harada A, Mori K (2008) ATF6 is a transcription factor specializing in the regulation of quality control proteins in the endoplasmic reticulum. Cell Struct Funct 33(1):75-89 doi:10.1247/csf.07044

Aggarwal RS (2014) What's fueling the biotech engine-2012 to 2013. Nat Biotechnol 32(1):32-39 doi:10.1038/nbt.2794

Alban A, David SO, Bjorkesten L, Andersson C, Sloge E, Lewis S, Currie I (2003) A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard. Proteomics 3(1):36-44 doi:10.1002/pmic.200390006

Anderson NL, Anderson NG (1998) Proteome and proteomics: new technologies, new concepts, and new words. Electrophoresis 19(11):1853-1861 doi:10.1002/elps.1150191103

Baik JY, Joo EJ, Kim YH, Lee GM (2008) Limitations to the comparative proteomic analysis of thrombopoietin producing Chinese hamster ovary cells treated with sodium butyrate. J Biotechnol 133(4):461-468 doi:10.1016/j.jbiotec.2007.11.008

Baik JY, Lee GM (2010) A DIGE approach for the assessment of differential expression of the CHO proteome under sodium butyrate addition: Effect of Bcl-x_L overexpression. Biotechnol Bioeng 105(2):358-367 doi:10.1002/bit.22534

Baik JY, Lee MS, An SR, Yoon SK, Joo EJ, Kim YH, Park HW, Lee GM (2006) Initial transcriptome and proteome analyses of low culture temperature-induced expression in CHO cells producing erythropoietin. Biotechnol Bioeng 93(2):361-371 doi:10.1002/bit.20717

Barnes LM, Bentley CM, Dickson AJ (2003) Stability of protein production from recombinant mammalian cells. Biotechnol Bioeng 81(6):631-639 doi:10.1002/bit.10517

Barron N, Kumar N, Sanchez N, Doolan P, Clarke C, Meleady P, O'Sullivan F, Clynes M (2011) Engineering CHO cell growth and recombinant protein productivity by overexpression of miR-7. J Biotechnol 151(2):204-211 doi:10.1016/j.jbiotec.2010.12.005

Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116(2):281-297 doi:10.1016/S0092-8674(04)00045-5

Baumann K, Dato L, Graf AB, Frascotti G, Dragosits M, Porro D, Mattanovich D, Ferrer P, Branduardi P (2011) The impact of oxygen on the transcriptome of recombinant *S. cerevisiae* and *P. pastoris* - a comparative analysis. BMC Genomics 12:218 doi:10.1186/1471-2164-12-218

Baycin-Hizal D, Tabb DL, Chaerkady R, Chen L, Lewis NE, Nagarajan H, Sarkaria V, Kumar A, Wolozny D, Colao J, Jacobson E, Tian Y, O'Meally RN, Krag SS, Cole RN, Palsson BO, Zhang H, Betenbaugh M (2012) Proteomic analysis of Chinese hamster ovary cells. J Proteome Res 11(11):5265-5276 doi:10.1021/pr300476w

Bebbington CR, Renner G, Thomson S, King D, Abrams D, Yarranton GT (1992) High-level expression of a recombinant antibody from myeloma cells using a glutamine synthetase gene as an amplifiable selectable marker. Biotechnology (N Y) 10(2):169-175 doi:10.1038/nbt0292-169

Beck A, Reichert JM (2012) Marketing approval of mogamulizumab: a triumph for glyco-engineering. MAbs 4(4):419-425 doi:10.4161/mabs.20996

Becker J, Hackl M, Rupp O, Jakobi T, Schneider J, Szczepanowski R, Bekel T, Borth N, Goesmann A, Grillari J, Kaltschmidt C, Noll T, Puhler A, Tauch A, Brinkrolf K (2011) Unraveling the Chinese hamster ovary cell line transcriptome by next-generation sequencing. J Biotechnol 156(3):227-235 doi:10.1016/j.jbiotec.2011.09.014

Benton T, Chen T, McEntee M, Fox B, King D, Crombie R, Thomas TC, Bebbington C (2002) The use of UCOE vectors in combination with a preadapted serum free, suspension cell line allows for rapid production of large quantities of protein. Cytotechnology 38(1-3):43-46 doi:10.1023/A:1021141712344

Birzele F, Schaub J, Rust W, Clemens C, Baum P, Kaufmann H, Weith A, Schulz TW, Hildebrandt T (2010) Into the unknown: expression profiling without genome sequence information in CHO by next generation sequencing. Nucleic Acids Res 38(12):3999-4010 doi:10.1093/nar/gkq116

Blaas L, Musteanu M, Eferl R, Bauer A, Casanova E (2009) Bacterial artificial chromosomes improve recombinant protein production in mammalian cells. BMC Biotechnol 9:3 doi:10.1186/1472-6750-9-3

Borth N, Mattanovich D, Kunert R, Katinger H (2005) Effect of increased expression of protein disulfide isomerase and heavy chain binding protein on antibody secretion in a recombinant CHO cell line. Biotechnol Prog 21(1):106-111 doi:10.1021/bp0498241

Bosques CJ, Collins BE, Meador JW, 3rd, Sarvaiya H, Murphy JL, Dellorusso G, Bulik DA, Hsu IH, Washburn N, Sipsey SF, Myette JR, Raman R, Shriver Z, Sasisekharan R, Venkataraman G (2010) Chinese hamster ovary cells can produce galactose-alpha-1,3-galactose antigens on proteins. Nat Biotechnol 28(11):1153-1156 doi:10.1038/nbt1110-1153

Bradley SA, Ouyang A, Purdie J, Smitka TA, Wang T, Kaerner A (2010) Fermentanomics: monitoring mammalian cell cultures with NMR spectroscopy. J Am Chem Soc 132(28):9531-9533 doi:10.1021/ja101962c

Brinkrolf K, Rupp O, Laux H, Kollin F, Ernst W, Linke B, Kofler R, Romand S, Hesse F, Budach WE, Galosy S, Müller D, Noll T, Wienberg J, Jostock T, Leonard M, Grillari J, Tauch A, Goesmann A, Helk B, Mott JE, Pühler A, Borth N (2013) Chinese hamster genome sequenced from sorted chromosomes. Nat Biotechnol 31(8):694-695 doi:10.1038/nbt.2645

Brown AJ, Sweeney B, Mainwaring DO, James DC (2014) Synthetic promoters for CHO cell engineering. Biotechnol Bioeng doi:10.1002/bit.25227

Cai H, Chen H, Yi T, Daimon CM, Boyle JP, Peers C, Maudsley S, Martin B (2013) VennPlex--a novel Venn diagram program for comparing and visualizing datasets with differentially regulated datapoints. PLOS ONE 8(1):e53388 doi:10.1371/journal.pone.0053388

Cao Y, Kimura S, Itoi T, Honda K, Ohtake H, Omasa T (2012) Construction of BAC-based physical map and analysis of chromosome rearrangement in Chinese hamster ovary cell lines. Biotechnol Bioeng 109(6):1357-1367 doi:10.1002/bit.24347

Carlage T, Hincapie M, Zang L, Lyubarskaya Y, Madden H, Mhatre R, Hancock WS (2009) Proteomic profiling of a high-producing Chinese hamster ovary cell culture. Anal Chem 81(17):7357-7362 doi:10.1021/ac900792z

Carlage T, Kshirsagar R, Zang L, Janakiraman V, Hincapie M, Lyubarskaya Y, Weiskopf A, Hancock WS (2012) Analysis of dynamic changes in the proteome of a Bcl-XL overexpressing Chinese hamster ovary cell culture during exponential and stationary phases. Biotechnol Prog 28(3):814-823 doi:10.1002/btpr.1534

Chen Z, Jin Y, Yu D, Wang A, Mahjabeen I, Wang C, Liu X, Zhou X (2012) Down-regulation of the microRNA-99 family members in head and neck squamous cell carcinoma. Oral Oncol 48(8):686-91 doi:10.1016/j.oraloncology.2012.02.020

Chiang GG, Sisk WP (2005) Bcl-x_L mediates increased production of humanized monoclonal antibodies in Chinese hamster ovary cells. Biotechnol Bioeng 91(7):779-792 doi:10.1002/bit.20551

Cho S, Jang I, Jun Y, Yoon S, Ko M, Kwon Y, Choi I, Chang H, Ryu D, Lee B, Kim VN, Kim W, Lee S (2013) MiRGator v3.0: a microRNA portal for deep sequencing, expression profiling and mRNA targeting. Nucleic Acids Res 41(Database issue):D252-D257 doi:10.1093/nar/gks1168

Chong WP, Goh LT, Reddy SG, Yusufi FN, Lee DY, Wong NS, Heng CK, Yap MG, Ho YS (2009) Metabolomics profiling of extracellular metabolites in recombinant Chinese Hamster Ovary fed-batch culture. Rapid Commun Mass Spectrom 23(23):3763-3771 doi:10.1002/rcm.4328

Chong WP, Reddy SG, Yusufi FN, Lee DY, Wong NS, Heng CK, Yap MG, Ho YS (2010) Metabolomics-driven approach for the improvement of Chinese hamster ovary cell growth: overexpression of malate dehydrogenase II. J Biotechnol 147(2):116-121 doi:10.1016/j.jbiotec.2010.03.018

Chong WP, Thng SH, Hiu AP, Lee DY, Chan EC, Ho YS (2012) LC-MS-based metabolic characterization of high monoclonal antibody-producing Chinese hamster ovary cells. Biotechnol Bioeng 109(12):3103-3111 doi:10.1002/bit.24580

Chong WP, Yusufi FN, Lee DY, Reddy SG, Wong NS, Heng CK, Yap MG, Ho YS (2011) Metabolomics-based identification of apoptosis-inducing metabolites in recombinant fed-batch CHO culture media. J Biotechnol 151(2):218-224 doi:10.1016/j.jbiotec.2010.12.010

Chung JY, Lim SW, Hong YJ, Hwang SO, Lee GM (2004) Effect of doxycycline-regulated calnexin and calreticulin expression on specific thrombopoietin productivity of recombinant Chinese hamster ovary cells. Biotechnol Bioeng 85(5):539-546 doi:10.1002/bit.10919

Clarke C, Henry M, Doolan P, Kelly S, Aherne S, Sanchez N, Kelly P, Kinsella P, Breen L, Madden SF, Zhang L, Leonard M, Clynes M, Meleady P, Barron N (2012) Integrated miRNA, mRNA and protein expression analysis reveals the role of post-transcriptional regulation in controlling CHO cell growth rate. BMC Genomics 13:656 doi:10.1186/1471-2164-13-656

Clincke MF, Molleryd C, Zhang Y, Lindskog E, Walsh K, Chotteau V (2013) Very high density of CHO cells in perfusion by ATF or TFF in WAVE bioreactor. Part I. Effect of the cell density on the process. Biotechnol Prog 29(3):754-767 doi:10.1002/btpr.1704

Cost GJ, Freyvert Y, Vafiadis A, Santiago Y, Miller JC, Rebar E, Collingwood TN, Snowden A, Gregory PD (2010) *BAK* and *BAX* deletion using zinc-finger nucleases yields apoptosis-resistant CHO cells. Biotechnol Bioeng 105(2):330-340 doi:10.1002/bit.22541

Cox J, Mann M (2007) Is proteomics the new genomics? Cell 130(3):395-398 doi:10.1016/j.cell.2007.07.032

Cui L, Zhou H, Zhao H, Zhou Y, Xu R, Xu X, Zheng L, Xue Z, Xia W, Zhang B, Ding T, Cao Y, Tian Z, Shi Q, He X (2012) MicroRNA-99a induces G1-phase cell cycle arrest and suppresses tumorigenicity in renal cell carcinoma. BMC Cancer 12:546 doi:10.1186/1471-2407-12-546

Curry S, Mandelkow H, Brick P, Franks N (1998) Crystal structure of human serum albumin complexed with fatty acid reveals an asymmetric distribution of binding sites. Nat Struct Biol 5(9):827-835 doi:10.1038/1869

Dai R, Li J, Liu Y, Yan D, Chen S, Duan C, Liu X, He T, Li H (2010) miR-221/222 suppression protects against endoplasmic reticulum stress-induced apoptosis via $p27^{Kip1}$- and MEK/ERK-mediated cell cycle regulation. Biol Chem 391(7):791-801 doi:10.1515/BC.2010.072

Dasgupta S, Navarrete AM, Bayry J, Delignat S, Wootla B, Andre S, Christophe O, Nascimbeni M, Jacquemin M, Martinez-Pomares L, Geijtenbeek TB, Moris A, Saint-Remy JM, Kazatchkine MD, Kaveri SV, Lacroix-Desmazes S (2007) A role for exposed mannosylations in presentation of human therapeutic self-proteins to CD4+ T lymphocytes. Proc Natl Acad Sci U S A 104(21):8965-8970 doi:10.1073/pnas.0702120104

Datta P, Linhardt RJ, Sharfstein ST (2013) An 'omics approach towards CHO cell engineering. Biotechnol Bioeng 110(5):1255-1271 doi:10.1002/bit.24841

Davis R, Schooley K, Rasmussen B, Thomas J, Reddy P (2000) Effect of PDI overexpression on recombinant protein secretion in CHO cells. Biotechnol Prog 16(5):736-743 doi:10.1021/bp000107q

De Jesus M, Wurm FM (2011) Manufacturing recombinant proteins in kg-ton quantities using animal cells in bioreactors. Eur J Pharm Biopharm 78(2):184-188 doi:10.1016/j.ejpb.2011.01.005

De Leon Gatti M, Wlaschin KF, Nissom PM, Yap M, Hu WS (2007) Comparative transcriptional analysis of mouse hybridoma and recombinant Chinese hamster ovary cells undergoing butyrate treatment. J Biosci Bioeng 103(1):82-91 doi:10.1263/jbb.103.82

Deaven LL, Petersen DF (1973) The chromosomes of CHO, an aneuploid Chinese hamster cell line: G-band, C-band, and autoradiographic analyses. Chromosoma 41(2):129-144 doi:10.1007/BF00319690

Dellago H, Preschitz-Kammerhofer B, Terlecki-Zaniewicz L, Schreiner C, Fortschegger K, Chang MW, Hackl M, Monteforte R, Kühnel H, Schosserer M, Gruber F, Tschachler E, Scheideler M, Grillari-Voglauer R, Grillari J, Wieser M (2013) High levels of oncomiR-21 contribute to the senescence-induced growth arrest in normal human cells and its knock-down increases the replicative lifespan. Aging Cell 12(3):446-458 doi:10.1111/acel.12069

Derouazi M, Martinet D, Besuchet Schmutz N, Flaction R, Wicht M, Bertschinger M, Hacker DL, Beckmann JS, Wurm FM (2006) Genetic characterization of CHO production host DG44 and derivative recombinant cell lines. Biochem Biophys Res Commun 340(4):1069-1077 doi:10.1016/j.bbrc.2005.12.111

Dharshanan S, Chong H, Hung CS, Zamrod Z, Kamal N (2011) Rapid automated selection of mammalian cell line secreting high level of humanized monoclonal antibody using Clone Pix FL system and the correlation between exterior median intensity and antibody productivity. Electronic Journal of Biotechnology 14(2) doi:10.2225/vol14-issue2-fulltext-7

Dietmair S, Hodson MP, Quek LE, Timmins NE, Chrysanthopoulos P, Jacob SS, Gray P, Nielsen LK (2012) Metabolite profiling of CHO cells with different growth characteristics. Biotechnol Bioeng 109(6):1404-1414 doi:10.1002/bit.24496

Dietmair S, Timmins NE, Gray PP, Nielsen LK, Kromer JO (2010) Towards quantitative metabolomics of mammalian cells: development of a metabolite extraction protocol. Anal Biochem 404(2):155-64 doi:10.1016/j.ab.2010.04.031

Doolan P, Meleady P, Barron N, Henry M, Gallagher R, Gammell P, Melville M, Sinacore M, McCarthy K, Leonard M, Charlebois T, Clynes M (2010) Microarray and proteomics expression profiling identifies several candidates, including the valosin-containing protein (VCP), involved in regulating high cellular growth rate in production CHO cell lines. Biotechnol Bioeng 106(1):42-56 doi:10.1002/bit.22670

Dragosits M, Stadlmann J, Albiol J, Baumann K, Maurer M, Gasser B, Sauer M, Altmann F, Ferrer P, Mattanovich D (2009) The effect of temperature on the proteome of recombinant *Pichia pastoris*. J Proteome Res 8(3):1380-1392 doi:10.1021/pr8007623

Dreesen IA, Fussenegger M (2011) Ectopic expression of human mTOR increases viability, robustness, cell size, proliferation, and antibody production of chinese hamster ovary cells. Biotechnol Bioeng 108(4):853-866 doi:10.1002/bit.22990

Druz A, Chu C, Majors B, Santuary R, Betenbaugh M, Shiloach J (2011) A novel microRNA mmu-miR-466h affects apoptosis regulation in mammalian cells. Biotechnol Bioeng 108(7):1651-1661 doi:10.1002/bit.23092

Druz A, Son YJ, Betenbaugh M, Shiloach J (2013) Stable inhibition of mmu-miR-466h-5p improves apoptosis resistance and protein production in CHO cells. Metab Eng 16:87-94 doi:10.1016/j.ymben.2012.12.004

Ernst W, Trummer E, Mead J, Bessant C, Strelec H, Katinger H, Hesse F (2006) Evaluation of a genomics platform for cross-species transcriptome analysis of recombinant CHO cells. Biotechnol J 1(6):639-650 doi:10.1002/biot.200600010

Ferrer-Miralles N, Domingo-Espín J, Corchero JL, Vázquez E, Villaverde A (2009) Microbial factories for recombinant pharmaceuticals. Microb Cell Fact 8:17 doi:10.1186/1475-2859-8-17

Fornari F, Gramantieri L, Ferracin M, Veronese A, Sabbioni S, Calin GA, Grazi GL, Giovannini C, Croce CM, Bolondi L, Negrini M (2008) MiR-221 controls CDKN1C/p57 and CDKN1B/p27 expression in human hepatocellular carcinoma. Oncogene 27(43):5651-5661 doi:10.1038/onc.2008.178

Fournier ML, Gilmore JM, Martin-Brown SA, Washburn MP (2007) Multidimensional separations-based shotgun proteomics. Chem Rev 107(8):3654-3686 doi:10.1021/cr068279a

Fujita S, Ito T, Mizutani T, Minoguchi S, Yamamichi N, Sakurai K, Iba H (2008) miR-21 Gene expression triggered by AP-1 is sustained through a double-negative feedback mechanism. J Mol Biol 378(3):492-504 doi:10.1016/j.jmb.2008.03.015

Gammell P, Barron N, Kumar N, Clynes M (2007) Initial identification of low temperature and culture stage induction of miRNA expression in suspension CHO-K1 cells. J Biotechnol 130(3):213-218 doi:10.1016/j.jbiotec.2007.04.020

Ge Y, Pan S, Guan D, Yin H, Fan Y, Liu J, Zhang S, Zhang H, Feng L, Wang Y, Xu R, Yin JQ (2013) MicroRNA-350 induces pathological heart hypertrophy by repressing both p38 and JNK pathways. Biochim Biophys Acta 1832(1):1-10 doi:10.1016/j.bbadis.2012.09.004

Gerstl MP, Hackl M, Graf AB, Borth N, Grillari J (2013) Prediction of transcribed PIWI-interacting RNAs from CHO RNAseq data. J Biotechnol 166(1-2):51-57 doi:10.1016/j.jbiotec.2013.04.010

Gilmore TD (2006) Introduction to NF-kappaB: players, pathways, perspectives. Oncogene 25(51):6680-6684 doi:10.1038/sj.onc.1209954

Girod PA, Zahn-Zabal M, Mermod N (2005) Use of the chicken lysozyme 5' matrix attachment region to generate high producer CHO cell lines. Biotechnol Bioeng 91(1):1-11 doi:10.1002/bit.20563

Gong J, Zhang JP, Li B, Zeng C, You K, Chen MX, Yuan Y, Zhuang SM (2013) MicroRNA-125b promotes apoptosis by regulating the expression of Mcl-1, Bcl-w and IL-6R. Oncogene 32(25):3071-3079 doi:10.1038/onc.2012.318

Grabherr MG, Pontiller J, Mauceli E, Ernst W, Baumann M, Biagi T, Swofford R, Russell P, Zody MC, Di Palma F, Lindblad-Toh K, Grabherr RM (2011) Exploiting nucleotide composition to engineer promoters. PLOS ONE 6(5):e20136 doi:10.1371/journal.pone.0020136

Grinna LS, Tschopp JF (1989) Size distribution and general structural features of N-linked oligosaccharides from the methylotrophic yeast, *Pichia pastoris*. Yeast 5(2):107-115 doi:10.1002/yea.320050206

Gustems M, Borst E, Benedict CA, Perez C, Messerle M, Ghazal P, Angulo A (2006) Regulation of the transcription and replication cycle of human cytomegalovirus is insensitive to genetic elimination of the cognate NF-kappaB binding sites in the enhancer. J Virol 80(19):9899-9904 doi:10.1128/JVI.00640-06

Hackl M, Borth N, Grillari J (2012a) The CHO miRNA transcriptome. In: Barron N (ed) MicroRNAs as Tools in Biopharmaceutical Production. Springer Netherlands, pp 49-64

Hackl M, Borth N, Grillari J (2012b) miRNAs--pathway engineering of CHO cell factories that avoids translational burdening. Trends Biotechnol 30(8):405-406 doi:10.1016/j.tibtech.2012.05.002

Hackl M, Jadhav V, Klanert G, Karbiener M, Scheideler M, Grillari J, Borth N (2014) Analysis of microRNA transcription and post-transcriptional processing by Dicer in the context of CHO cell proliferation. J Biotechnol doi:10.1016/j.jbiotec.2013.12.018

Hackl M, Jakobi T, Blom J, Doppmeier D, Brinkrolf K, Szczepanowski R, Bernhart SH, Höner zu Siederdissen C, Hernandez Bort JA, Wieser M, Kunert R, Jeffs S, Hofacker IL, Goesmann A, Pühler A, Borth N, Grillari J (2011) Next-generation sequencing of the Chinese hamster ovary microRNA transcriptome: Identification, annotation and profiling of microRNAs as targets for cellular engineering. J Biotechnol 153(1-2):62-75 doi:10.1016/j.jbiotec.2011.02.011

Hamilton SR, Davidson RC, Sethuraman N, Nett JH, Jiang Y, Rios S, Bobrowicz P, Stadheim TA, Li H, Choi BK, Hopkins D, Wischnewski H, Roser J, Mitchell T, Strawbridge RR, Hoopes J, Wildt S, Gerngross TU (2006) Humanization of yeast to produce complex terminally sialylated glycoproteins. Science 313(5792):1441-1443 doi:10.1126/science.1130256

Hammond S, Kaplarevic M, Borth N, Betenbaugh MJ, Lee KH (2012a) Chinese hamster genome database: an online resource for the CHO community at www.CHOgenome.org. Biotechnol Bioeng 109(6):1353-1356 doi:10.1002/bit.24374

Hammond S, Swanberg JC, Polson SW, Lee KH (2012b) Profiling conserved microRNA expression in recombinant CHO cell lines using Illumina sequencing. Biotechnol Bioeng 109(6):1371-1375 doi:10.1002/bit.24415

Han J, Back SH, Hur J, Lin YH, Gildersleeve R, Shan J, Yuan CL, Krokowski D, Wang S, Hatzoglou M, Kilberg MS, Sartor MA, Kaufman RJ (2013) ER-stress-induced transcriptional regulation increases protein synthesis leading to cell death. Nat Cell Biol 15(5):481-490 doi:10.1038/ncb2738

Hayduk EJ, Lee KH (2005) Cytochalasin D can improve heterologous protein productivity in adherent Chinese hamster ovary cells. Biotechnol Bioeng 90(3):354-364 doi:10.1002/bit.20438

Hayter PM, Curling EM, Baines AJ, Jenkins N, Salmon I, Strange PG, Tong JM, Bull AT (1992) Glucose-limited chemostat culture of Chinese hamster ovary cells producing recombinant human interferon-gamma. Biotechnol Bioeng 39(3):327-335 doi:10.1002/bit.260390311

He J, Li Y, Yang X, He X, Zhang H, He J, Zhang L (2013) The feedback regulation of PI3K-miR-19a, and MAPK-miR-23b/27b in endothelial cells under shear stress. Molecules 18(1):1-13 doi:10.3390/molecules18010001

He L, Hannon GJ (2004) MicroRNAs: small RNAs with a big role in gene regulation. Nat Rev Genet 5(7):522-531 doi:10.1038/nrg1379

Hernandez Bort JA, Hackl M, Höflmayer H, Jadhav V, Harreither E, Kumar N, Ernst W, Grillari J, Borth N (2012) Dynamic mRNA and miRNA profiling of CHO-K1 suspension cell cultures. Biotechnol J 7(4):500-515 doi:10.1002/biot.201100143

Hernandez Bort JA, Shanmukam V, Pabst M, Windwarder M, Neumann L, Alchalabi A, Krebiehl G, Koellensperger G, Hann S, Sonntag D, Altmann F, Heel C, Borth N (2014) Reduced quenching and extraction time for mammalian cells using filtration and syringe extraction. J Biotechnol 182-183:97-103 doi:10.1016/j.jbiotec.2014.04.014

Hobert O (2008) Gene regulation by transcription factors and microRNAs. Science 319(5871):1785-1786 doi:10.1126/science.1151651

Hori O, Miyazaki M, Tamatani T, Ozawa K, Takano K, Okabe M, Ikawa M, Hartmann E, Mai P, Stern DM, Kitao Y, Ogawa S (2006) Deletion of SERP1/RAMP4, a component of the endoplasmic reticulum (ER) translocation sites, leads to ER stress. Mol Cell Biol 26(11):4257-4267 doi:10.1128/MCB.02055-05

Hoskisson PA, Hobbs G (2005) Continuous culture--making a comeback? Microbiology 151(Pt 10):3153-3159 doi:10.1099/mic.0.27924-0

Hsu SD, Tseng YT, Shrestha S, Lin YL, Khaleel A, Chou CH, Chu CF, Huang HY, Lin CM, Ho SY, Jian TY, Lin FM, Chang TH, Weng SL, Liao KW, Liao IE, Liu CC, Huang HD (2014) miRTarBase update 2014: an information resource for experimentally validated miRNA-target interactions. Nucleic Acids Res 42(Database issue):D78-D85 doi:10.1093/nar/gkt1266

Hwang SO, Chung JY, Lee GM (2003) Effect of doxycycline-regulated ERp57 expression on specific thrombopoietin productivity of recombinant CHO cells. Biotechnol Prog 19(1):179-184 doi:10.1021/bp025578m

Ibrahim SA, Yip GW, Stock C, Pan JW, Neubauer C, Poeter M, Pupjalis D, Koo CY, Kelsch R, Schüle R, Rescher U, Kiesel L, Götte M (2012) Targeting of syndecan-1 by microRNA miR-10b promotes breast cancer cell motility and invasiveness via a Rho-GTPase- and E-cadherin-dependent mechanism. Int J Cancer 131(6):E884-E896 doi:10.1002/ijc.27629

Jadhav V, Hackl M, Bort JA, Wieser M, Harreither E, Kunert R, Borth N, Grillari J (2012) A screening method to assess biological effects of microRNA overexpression in Chinese hamster ovary cells. Biotechnol Bioeng 109(6):1376-1385 doi:10.1002/bit.24490

Jadhav V, Hackl M, Druz A, Shridhar S, Chung CY, Heffner KM, Kreil DP, Betenbaugh M, Shiloach J, Barron N, Grillari J, Borth N (2013) CHO microRNA engineering is growing up: recent successes and future challenges. Biotechnol Adv 31(8):1501-1513 doi:10.1016/j.biotechadv.2013.07.007

Jadhav V, Hackl M, Klanert G, Bort JA, Kunert R, Borth N, Grillari J (2014) Stable overexpression of miR-17 enhances recombinant protein production of CHO cells. J Biotechnol 175:38-44 doi:10.1016/j.jbiotec.2014.01.032

Jakymiw A, Patel RS, Deming N, Bhattacharyya I, Shah P, Lamont RJ, Stewart CM, Cohen DM, Chan EK (2010) Overexpression of dicer as a result of reduced let-7 MicroRNA levels contributes to increased cell proliferation of oral cancer cells. Genes Chromosomes Cancer 49(6):549-559 doi:10.1002/gcc.20765

Jaluria P, Betenbaugh M, Konstantopoulos K, Shiloach J (2007) Enhancement of cell proliferation in various mammalian cell lines by gene insertion of a cyclin-dependent kinase homolog. BMC Biotechnol 7:71 doi:10.1186/1472-6750-7-71

Jayapal KR, Wlaschin KF, Hu WS, Yap MGS (2007) Recombinant protein therapeutics from CHO cells - 20 years and counting. Chemical Engineering Progress 103(10):40-47

Jefferis R (2009) Recombinant antibody therapeutics: the impact of glycosylation on mechanisms of action. Trends in Pharmacological Sciences 30(7):356-362 doi:DOI 10.1016/j.tips.2009.04.007

Jensen D, Schekman R (2011) COPII-mediated vesicle formation at a glance. J Cell Sci 124(Pt 1):1-4 doi:10.1242/jcs.069773

Jensen ON (2004) Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. Curr Opin Chem Biol 8(1):33-41 doi:10.1016/j.cbpa.2003.12.009

Jeong DW, Cho IT, Kim TS, Bae GW, Kim IH, Kim IY (2006) Effects of lactate dehydrogenase suppression and glycerol-3-phosphate dehydrogenase overexpression on cellular metabolism. Mol Cell Biochem 284(1-2):1-8 doi:10.1007/s11010-005-9004-7

Johnson KC, Jacob NM, Nissom PM, Hackl M, Lee LH, Yap M, Hu WS (2011) Conserved microRNAs in Chinese hamster ovary cell lines. Biotechnol Bioeng 108(2):475-480 doi:10.1002/bit.22940

Kang S, Ren D, Xiao G, Daris K, Buck L, Enyenihi AA, Zubarev R, Bondarenko PV, Deshpande R (2014) Cell line profiling to improve monoclonal antibody production. Biotechnol Bioeng 111(4):748-760 doi:10.1002/bit.25141

Kantardjieff A, Jacob NM, Yee JC, Epstein E, Kok YJ, Philp R, Betenbaugh M, Hu WS (2010) Transcriptome and proteome analysis of Chinese hamster ovary cells under low temperature and butyrate treatment. J Biotechnol 145(2):143-159 doi:10.1016/j.jbiotec.2009.09.008

Kantardjieff A, Nissom PM, Chuah SH, Yusufi F, Jacob NM, Mulukutla BC, Yap M, Hu WS (2009) Developing genomic platforms for Chinese hamster ovary cells. Biotechnol Adv 27(6):1028-1035 doi:10.1016/j.biotechadv.2009.05.023

Kao FT, Puck TT (1967) Genetics of somatic mammalian cells. IV. Properties of Chinese hamster cell mutants with respect to the requirement for proline. Genetics 55(3):513-524

Kaufman RJ, Sharp PA (1982) Amplification and expression of sequences cotransfected with a modular dihydrofolate reductase complementary DNA gene. J Mol Biol 159(4):601-621 doi:10.1016/0022-2836(82)90103-6

Kaufmann H, Mazur X, Fussenegger M, Bailey JE (1999) Influence of low temperature on productivity, proteome and protein phosphorylation of CHO cells. Biotechnol Bioeng 63(5):573-582

Kennard ML, Goosney DL, Monteith D, Zhang L, Moffat M, Fischer D, Mott J (2009) The generation of stable, high MAb expressing CHO cell lines based on the artificial chromosome expression (ACE) technology. Biotechnol Bioeng 104(3):540-553 doi:10.1002/bit.22406

Kim JD, Yoon Y, Hwang HY, Park JS, Yu S, Lee J, Baek K, Yoon J (2005) Efficient selection of stable chinese hamster ovary (CHO) cell lines for expression of recombinant proteins by using human interferon beta SAR element. Biotechnol Prog 21(3):933-937 doi:10.1021/bp049598v

Kim JM, Kim JS, Park DH, Kang HS, Yoon J, Baek K, Yoon Y (2004) Improved recombinant gene expression in CHO cells using matrix attachment regions. J Biotechnol 107(2):95-105 doi:10.1016/j.jbiotec.2003.09.015

Kim JY, Kim YG, Lee GM (2012) CHO cells in biotechnology for production of recombinant proteins: current state and further potential. Appl Microbiol Biotechnol 93(3):917-930 doi:10.1007/s00253-011-3758-5

Kim NS, Lee GM (2001) Overexpression of *bcl-2* inhibits sodium butyrate-induced apoptosis in Chinese hamster ovary cells resulting in enhanced humanized antibody production. Biotechnol Bioeng 71(3):184-193 doi:10.1002/1097-0290(2000)71:3<184::AID-BIT1008>3.0.CO;2-W

Kim NS, Lee GM (2002) Inhibition of sodium butyrate-induced apoptosis in recombinant Chinese hamster ovary cells by constitutively expressing antisense RNA of caspase-3. Biotechnol Bioeng 78(2):217-228 doi:10.1002/bit.10191

Kito M, Itami S, Fukano Y, Yamana K, Shibui T (2002) Construction of engineered CHO strains for high-level production of recombinant proteins. Appl Microbiol Biotechnol 60(4):442-448 doi:10.1007/s00253-002-1134-1

Kobayashi K, Kuwae S, Ohya T, Ohda T, Ohyama M, Ohi H, Tomomitsu K, Ohmura T (2000) High-level expression of recombinant human serum albumin from the methylotrophic yeast *Pichia pastoris* with minimal protease production and activation. J Biosci Bioeng 89(1):55-61 doi:10.1016/S1389-1723(00)88050-0

Koh TC, Lee YY, Chang SQ, Nissom PM (2009) Identification and expression analysis of miRNAs during batch culture of HEK-293 cells. J Biotechnol 140(3-4):149-155 doi:10.1016/j.jbiotec.2009.01.021

Kozomara A, Griffiths-Jones S (2014) miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res 42(Database issue):D68-D73 doi:10.1093/nar/gkt1181

Krichevsky AM, Gabriely G (2009) miR-21: a small multi-faceted RNA. J Cell Mol Med 13(1):39-53 doi:10.1111/j.1582-4934.2008.00556.x

Krol J, Loedige I, Filipowicz W (2010) The widespread regulation of microRNA biogenesis, function and decay. Nat Rev Genet 11(9):597-610 doi:10.1038/nrg2843

Ku SC, Ng DT, Yap MG, Chao SH (2008) Effects of overexpression of X-box binding protein 1 on recombinant protein production in Chinese hamster ovary and NS0 myeloma cells. Biotechnol Bioeng 99(1):155-164 doi:10.1002/bit.21562

Kumar N, Gammell P, Meleady P, Henry M, Clynes M (2008) Differential protein expression following low temperature culture of suspension CHO-K1 cells. BMC Biotechnol 8:42 doi:10.1186/1472-6750-8-42

Kuystermans D, Al-Rubeai M (2009) cMyc increases cell number through uncoupling of cell division from cell size in CHO cells. BMC Biotechnol 9:76 doi:10.1186/1472-6750-9-76

Lai T, Yang Y, Ng SK (2013) Advances in Mammalian cell line development technologies for recombinant protein production. Pharmaceuticals (Basel) 6(5):579-603 doi:10.3390/ph6050579

Lao MS, Toth D (1997) Effects of ammonium and lactate on growth and metabolism of a recombinant Chinese hamster ovary cell culture. Biotechnology Progress 13(5):688-691 doi:10.1021/Bp9602360

Le Fourn V, Girod PA, Buceta M, Regamey A, Mermod N (2014) CHO cell engineering to prevent polypeptide aggregation and improve therapeutic protein secretion. Metab Eng 21:91-102 doi:10.1016/j.ymben.2012.12.003

Lee MS, Kim KW, Kim YH, Lee GM (2003) Proteome analysis of antibody-expressing CHO cells in response to hyperosmotic pressure. Biotechnol Prog 19(6):1734-1741 doi:10.1021/bp034093a

Lee SY (1996) High cell-density culture of *Escherichia coli*. Trends Biotechnol 14(3):98-105 doi:10.1016/0167-7799(96)80930-9

Lewis NE, Liu X, Li Y, Nagarajan H, Yerganian G, O'Brien E, Bordbar A, Roth AM, Rosenbloom J, Bian C, Xie M, Chen W, Li N, Baycin-Hizal D, Latif H, Forster J, Betenbaugh MJ, Famili I, Xu X, Wang J, Palsson BO (2013) Genomic landscapes of Chinese hamster ovary cell lines as revealed by the *Cricetulus griseus* draft genome. Nat Biotechnol 31(8):759-765 doi:10.1038/nbt.2624

Li JH, Huang Z, Sun XM, Yang PY, Zhang YX (2006) Understanding the enhanced effect of dimethyl sulfoxide on hepatitis B surface antigen expression in the culture of Chinese hamster ovary cells on the basis of proteome analysis. Enzyme and Microbial Technology 38(3-4):372-380 doi:10.1016/j.enzmictec.2005.05.016

Li Y, Gao L, Luo X, Wang L, Gao X, Wang W, Sun J, Dou L, Li J, Xu C, Wang L, Zhou M, Jiang M, Zhou J, Caligiuri MA, Nervi C, Bloomfield CD, Marcucci G, Yu L (2013) Epigenetic silencing of microRNA-193a contributes to leukemogenesis in t(8;21) acute myeloid leukemia by activating the PTEN/PI3K signal pathway. Blood 121(3):499-509 doi:10.1182/blood-2012-07-444729

Li Z, Nabel GJ (1997) A new member of the I kappaB protein family, I kappaB epsilon, inhibits RelA (p65)-mediated NF-kappaB transcription. Mol Cell Biol 17(10):6184-6190

Liang B, Tikhanovich I, Nasheuer HP, Folk WR (2012) Stimulation of BK virus DNA replication by NFI family transcription factors. J Virol 86(6):3264-3275 doi:10.1128/JVI.06369-11

Lin J, Teo S, Lam DH, Jeyaseelan K, Wang S (2012) MicroRNA-10b pleiotropically regulates invasion, angiogenicity and apoptosis of tumor cells resembling mesenchymal subtype of glioblastoma multiforme. Cell Death Dis 3:e398 doi:10.1038/cddis.2012.134

Lin N, Davis A, Bahr S, Borgschulte T, Achtien K, Kayser K (2011) Profiling highly conserved microrna expression in recombinant IgG-producing and parental Chinese hamster ovary cells. Biotechnol Prog 27(4):1163-1171 doi:10.1002/btpr.556

Liu J, Ku SC, Lee J, Young TM, Pe'ery T, Mathews MB, Chao SH (2009) The 3'UTR of HIC mRNA improves the production of recombinant proteins in Chinese hamster ovary cells. J Biotechnol 139(2):152-155 doi:10.1016/j.jbiotec.2008.10.008

Lu TP, Lee CY, Tsai MH, Chiu YC, Hsiao CK, Lai LC, Chuang EY (2012) miRSystem: an integrated system for characterizing enriched functions and pathways of microRNA targets. PLOS ONE 7(8):e42390 doi:10.1371/journal.pone.0042390

Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM (2005) Heterologous protein production using the *Pichia pastoris* expression system. Yeast 22(4):249-270 doi:10.1002/yea.1208

Maccani A, Ernst W, Grabherr R (2013) Whole genome sequencing improves estimation of nuclear DNA content of Chinese hamster ovary cells. Cytometry Part A 83(10):893-895 doi:10.1002/Cyto.A.22331

Maccani A, Landes N, Stadlmayr G, Maresch D, Leitner C, Maurer M, Gasser B, Ernst W, Kunert R, Mattanovich D (2014) *Pichia pastoris* secretes recombinant proteins less efficiently than Chinese hamster ovary cells but allows higher space-time yields for less complex proteins. Biotechnol J 9(4):526-537 doi:10.1002/biot.201300305

Mader A, Prewein B, Zboray K, Casanova E, Kunert R (2013) Exploration of BAC versus plasmid expression vectors in recombinant CHO cells. Appl Microbiol Biotechnol 97(9):4049-4054 doi:10.1007/s00253-012-4498-x

Majors BS, Arden N, Oyler GA, Chiang GG, Pederson NE, Betenbaugh MJ (2008) E2F-1 overexpression increases viable cell density in batch cultures of Chinese hamster ovary cells. J Biotechnol 138(3-4):103-106 doi:10.1016/j.jbiotec.2008.08.003

Majors BS, Betenbaugh MJ, Pederson NE, Chiang GG (2009) Mcl-1 overexpression leads to higher viabilities and increased production of humanized monoclonal antibody in Chinese hamster ovary cells. Biotechnol Prog 25(4):1161-1168 doi:10.1002/btpr.192

Makrides SC (1999) Components of vectors for gene transfer and expression in mammalian cells. Protein Expr Purif 17(2):183-202 doi:10.1006/prep.1999.1137

Mattanovich D, Branduardi P, Dato L, Gasser B, Sauer M, Porro D (2012) Recombinant protein production in yeasts. Methods Mol Biol 824:329-358 doi:10.1007/978-1-61779-433-9_17

Mattanovich D, Graf A, Stadlmann J, Dragosits M, Redl A, Maurer M, Kleinheinz M, Sauer M, Altmann F, Gasser B (2009) Genome, secretome and glucose transport highlight unique features of the protein production host *Pichia pastoris*. Microb Cell Fact 8:29 doi:10.1186/1475-2859-8-29

Meents H, Enenkel B, Eppenberger HM, Werner RG, Fussenegger M (2002a) Impact of coexpression and coamplification of sICAM and antiapoptosis determinants *bcl-2/bcl-x_L* on productivity, cell survival, and mitochondria number in CHO-DG44 grown in suspension and serum-free media. Biotechnol Bioeng 80(6):706-716 doi:10.1002/bit.10449

Meents H, Enenkel B, Werner RG, Fussenegger M (2002b) $p27^{Kip1}$-mediated controlled proliferation technology increases constitutive sICAM production in CHO-DUKX adapted for growth in suspension and serum-free media. Biotechnol Bioeng 79(6):619-627 doi:10.1002/bit.10322

Meleady P, Gallagher M, Clarke C, Henry M, Sanchez N, Barron N, Clynes M (2012a) Impact of miR-7 over-expression on the proteome of Chinese hamster ovary cells. J Biotechnol 160(3-4):251-262 doi:10.1016/j.jbiotec.2012.03.002

Meleady P, Henry M, Gammell P, Doolan P, Sinacore M, Melville M, Francullo L, Leonard M, Charlebois T, Clynes M (2008) Proteomic profiling of CHO cells with enhanced rhBMP-2 productivity following co-expression of PACEsol. Proteomics 8(13):2611-2624 doi:10.1002/pmic.200700854

Meleady P, Hoffrogge R, Henry M, Rupp O, Bort JH, Clarke C, Brinkrolf K, Kelly S, Muller B, Doolan P, Hackl M, Beckmann TF, Noll T, Grillari J, Barron N, Puhler A, Clynes M, Borth N (2012b) Utilization and evaluation of CHO-specific sequence databases for mass spectrometry based proteomics. Biotechnol Bioeng 109(6):1386-1394 doi:10.1002/bit.24476

Melville M, Doolan P, Mounts W, Barron N, Hann L, Leonard M, Clynes M, Charlebois T (2011) Development and characterization of a Chinese hamster ovary cell-specific oligonucleotide microarray. Biotechnol Lett 33(9):1773-1779 doi:10.1007/s10529-011-0628-2

Mohan C, Kim YG, Koo J, Lee GM (2008) Assessment of cell engineering strategies for improved therapeutic protein production in CHO cells. Biotechnol J 3(5):624-630 doi:10.1002/biot.200700249

Mohan C, Park SH, Chung JY, Lee GM (2007) Effect of doxycycline-regulated protein disulfide isomerase expression on the specific productivity of recombinant CHO cells: thrombopoietin and antibody. Biotechnol Bioeng 98(3):611-615 doi:10.1002/bit.21453

Müller D, Katinger H, Grillari J (2008) MicroRNAs as targets for engineering of CHO cell factories. Trends Biotechnol 26(7):359-365 doi:10.1016/j.tibtech.2008.03.010

Ng SK, Wang DI, Yap MG (2007) Application of destabilizing sequences on selection marker for improved recombinant protein productivity in CHO-DG44. Metab Eng 9(3):304-316 doi:10.1016/j.ymben.2007.01.001

Ni Y, Chen R (2009) Extracellular recombinant protein production from *Escherichia coli*. Biotechnol Lett 31(11):1661-1670 doi:10.1007/s10529-009-0077-3

Nissom PM, Sanny A, Kok YJ, Hiang YT, Chuah SH, Shing TK, Lee YY, Wong KT, Hu WS, Sim MY, Philp R (2006) Transcriptome and proteome profiling to understanding the biology of high productivity CHO cells. Mol Biotechnol 34(2):125-140 doi:10.1385/MB:34:2:125

Niwa H, Yamamura K, Miyazaki J (1991) Efficient selection for high-expression transfectants with a novel eukaryotic vector. Gene 108(2):193-199 doi:10.1016/0378-1119(91)90434-D

Ohya T, Hayashi T, Kiyama E, Nishii H, Miki H, Kobayashi K, Honda K, Omasa T, Ohtake H (2008) Improved production of recombinant human antithrombin III in Chinese hamster ovary cells by ATF4 overexpression. Biotechnol Bioeng 100(2):317-324 doi:10.1002/bit.21758

Okada T, Yoshida H, Akazawa R, Negishi M, Mori K (2002) Distinct roles of activating transcription factor 6 (ATF6) and double-stranded RNA-activated protein kinase-like endoplasmic reticulum kinase (PERK) in transcription during the mammalian unfolded protein response. Biochem J 366(Pt 2):585-594 doi:10.1042/BJ20020391

Omasa T, Takami T, Ohya T, Kiyama E, Hayashi T, Nishii H, Miki H, Kobayashi K, Honda K, Ohtake H (2008) Overexpression of GADD34 enhances production of recombinant human antithrombin III in Chinese hamster ovary cells. J Biosci Bioeng 106(6):568-573 doi:10.1263/jbb.106.568

Panda AK (2003) Bioprocessing of therapeutic proteins from the inclusion bodies of *Escherichia coli*. Adv Biochem Eng Biotechnol 85:43-93 doi:10.1007/3-540-36466-8_3

Pandhal J, Wright P (2010) N-Linked glycoengineering for human therapeutic proteins in bacteria. Biotechnology Letters 32(9):1189-1198 doi:10.1007/s10529-010-0289-6

Park H, Kim IH, Kim IY, Kim KH, Kim HJ (2000) Expression of carbamoyl phosphate synthetase I and ornithine transcarbamoylase genes in Chinese hamster ovary *dhfr*-cells decreases accumulation of ammonium ion in culture media. J Biotechnol 81(2-3):129-140 doi:10.1016/S0168-1656(00)00282-0

Patti GJ, Yanes O, Siuzdak G (2012) Innovation: Metabolomics: the apogee of the omics trilogy. Nat Rev Mol Cell Biol 13(4):263-269 doi:10.1038/nrm3314

Peng RW, Abellan E, Fussenegger M (2011) Differential effect of exocytic SNAREs on the production of recombinant proteins in mammalian cells. Biotechnol Bioeng 108(3):611-620 doi:10.1002/bit.22986

Peng RW, Fussenegger M (2009) Molecular engineering of exocytic vesicle traffic enhances the productivity of Chinese hamster ovary cells. Biotechnol Bioeng 102(4):1170-1181 doi:10.1002/bit.22141

Potgieter TI, Cukan M, Drummond JE, Houston-Cummings NR, Jiang YW, Li F, Lynaugh H, Mallem M, McKelvey TW, Mitchell T, Nylen A, Rittenhour A, Stadheim TA, Zha DX, d'Anjou M (2009) Production of monoclonal antibodies by glycoengineered *Pichia pastoris*. Journal of Biotechnology 139(4):318-325 doi:DOI 10.1016/j.jbiotec.2008.12.015

Puck TT, Kao FT (1967) Genetics of somatic mammalian cells. V. Treatment with 5-bromodeoxyuridine and visible light for isolation of nutritionally deficient mutants. Proc Natl Acad Sci U S A 58(3):1227-1234

Qin X, Wang X, Wang Y, Tang Z, Cui Q, Xi J, Li YS, Chien S, Wang N (2010) MicroRNA-19a mediates the suppressive effect of laminar flow on cyclin D1 expression in human umbilical vein endothelial cells. Proc Natl Acad Sci U S A 107(7):3240-3244 doi:10.1073/pnas.0914882107

Rabilloud T, Chevallet M, Luche S, Lelong C (2010) Two-dimensional gel electrophoresis in proteomics: Past, present and future. J Proteomics 73(11):2064-2077 doi:10.1016/j.jprot.2010.05.016

Reichert JM (2012) Marketed therapeutic antibodies compendium. MAbs 4(3):413-415 doi:10.4161/mabs.19931

Rickwood S, Kleinrock M, Núñez-Gaviria M (2013) The Global Use of Medicines: Outlook through 2017. IMS Institute for Healthcare Informatics

Roush S, Slack FJ (2008) The let-7 family of microRNAs. Trends Cell Biol 18(10):505-516 doi:10.1016/j.tcb.2008.07.007

Running Deer J, Allison DS (2004) High-level expression of proteins in mammalian cells using transcription regulatory sequences from the Chinese hamster EF-1alpha gene. Biotechnol Prog 20(3):880-889 doi:10.1021/bp034383r

Rupp O, Becker J, Brinkrolf K, Timmermann C, Borth N, Puhler A, Noll T, Goesmann A (2014) Construction of a public CHO cell line transcript database using versatile bioinformatics analysis pipelines. PLOS ONE 9(1):e85568 doi:10.1371/journal.pone.0085568

Sauerwald TM, Betenbaugh MJ, Oyler GA (2002) Inhibiting apoptosis in mammalian cell culture using the caspase inhibitor XIAP and deletion mutants. Biotechnol Bioeng 77(6):704-716 doi:10.1002/bit.10154

Sautter K, Enenkel B (2005) Selection of high-producing CHO cells using NPT selection marker with reduced enzyme activity. Biotechnol Bioeng 89(5):530-538 doi:10.1002/bit.20374

Schröder K, Martoglio B, Hofmann M, Hölscher C, Hartmann E, Prehn S, Rapoport TA, Dobberstein B (1999) Control of glycosylation of MHC class II-associated invariant chain by translocon-associated RAMP4. EMBO J 18(17):4804-4815 doi:10.1093/emboj/18.17.4804

Sellick CA, Croxford AS, Maqsood AR, Stephens G, Westerhoff HV, Goodacre R, Dickson AJ (2011a) Metabolite profiling of recombinant CHO cells: designing tailored feeding regimes that enhance recombinant antibody production. Biotechnol Bioeng 108(12):3025-3031 doi:10.1002/bit.23269

Sellick CA, Hansen R, Stephens GM, Goodacre R, Dickson AJ (2011b) Metabolite extraction from suspension-cultured mammalian cells for global metabolite profiling. Nat Protoc 6(8):1241-1249 doi:10.1038/nprot.2011.366

Seth G, Charaniya S, Wlaschin KF, Hu WS (2007) In pursuit of a super producer-alternative paths to high producing recombinant mammalian cells. Curr Opin Biotechnol 18(6):557-564 doi:10.1016/j.copbio.2007.10.012

Shaffer AL, Shapiro-Shelef M, Iwakoshi NN, Lee AH, Qian SB, Zhao H, Yu X, Yang L, Tan BK, Rosenwald A, Hurt EM, Petroulakis E, Sonenberg N, Yewdell JW, Calame K, Glimcher LH, Staudt LM (2004) XBP1, downstream of Blimp-1, expands the secretory apparatus and other organelles, and increases protein synthesis in plasma cell differentiation. Immunity 21(1):81-93 doi:10.1016/j.immuni.2004.06.010

Shen D, Kiehl TR, Khattak SF, Li ZJ, He A, Kayne PS, Patel V, Neuhaus IM, Sharfstein ST (2010) Transcriptomic responses to sodium chloride-induced osmotic stress: a study of industrial fed-batch CHO cell cultures. Biotechnol Prog 26(4):1104-1115 doi:10.1002/btpr.398

Shibahara K, Asano M, Ishida Y, Aoki T, Koike T, Honjo T (1995) Isolation of a novel mouse gene MA-3 that is induced upon programmed cell death. Gene 166(2):297-301 doi:10.1016/0378-1119(95)00607-9

Shoulders MD, Ryno LM, Genereux JC, Moresco JJ, Tu PG, Wu C, Yates JR, 3rd, Su AI, Kelly JW, Wiseman RL (2013) Stress-independent activation of XBP1s and/or ATF6 reveals three functionally diverse ER proteostasis environments. Cell Rep 3(4):1279-1292 doi:10.1016/j.celrep.2013.03.024

Sinha J, Plantz BA, Inan M, Meagher MM (2005) Causes of proteolytic degradation of secreted recombinant proteins produced in methylotrophic yeast *Pichia pastoris*: case study with recombinant ovine interferon-tau. Biotechnol Bioeng 89(1):102-112 doi:10.1002/bit.20318

Sola RJ, Griebenow K (2009) Effects of glycosylation on the stability of protein pharmaceuticals. J Pharm Sci 98(4):1223-1245 doi:10.1002/jps.21504

Stinski MF, Isomura H (2008) Role of the cytomegalovirus major immediate early enhancer in acute infection and reactivation from latency. Med Microbiol Immunol 197(2):223-231 doi:10.1007/s00430-007-0069-7

Sun J, Chen Z, Tan X, Zhou F, Tan F, Gao Y, Sun N, Xu X, Shao K, He J (2013) MicroRNA-99a/100 promotes apoptosis by targeting mTOR in human esophageal squamous cell carcinoma. Med Oncol 30(1):411 doi:10.1007/s12032-012-0411-9

Sung YH, Lee JS, Park SH, Koo J, Lee GM (2007) Influence of co-down-regulation of caspase-3 and caspase-7 by siRNAs on sodium butyrate-induced apoptotic cell death of Chinese hamster ovary cells producing thrombopoietin. Metab Eng 9(5-6):452-464 doi:10.1016/j.ymben.2007.08.001

Suvasini R, Shruti B, Thota B, Shinde SV, Friedmann-Morvinski D, Nawaz Z, Prasanna KV, Thennarasu K, Hegde AS, Arivazhagan A, Chandramouli BA, Santosh V, Somasundaram K (2011) Insulin growth factor-2 binding protein 3 (IGF2BP3) is a glioblastoma-specific marker that activates phosphatidylinositol 3-kinase/mitogen-activated protein kinase (PI3K/MAPK) pathways by modulating IGF-2. J Biol Chem 286(29):25882-90 doi:10.1074/jbc.M110.178012

Suzuki C, Garces RG, Edmonds KA, Hiller S, Hyberts SG, Marintchev A, Wagner G (2008) PDCD4 inhibits translation initiation by binding to eIF4A using both its MA3 domains. Proc Natl Acad Sci U S A 105(9):3274-3279 doi:10.1073/pnas.0712235105

Szatmari T, Dobra K (2013) The role of syndecan-1 in cellular signaling and its effects on heparan sulfate biosynthesis in mesenchymal tumors. Front Oncol 3:310 doi:10.3389/fonc.2013.00310

Tait AS, Hogwood CE, Smales CM, Bracewell DG (2012) Host cell protein dynamics in the supernatant of a mAb producing CHO cell line. Biotechnol Bioeng 109(4):971-982 doi:10.1002/bit.24383

Tan HK, Lee MM, Yap MG, Wang DI (2008) Overexpression of cold-inducible RNA-binding protein increases interferon-gamma production in Chinese-hamster ovary cells. Biotechnol Appl Biochem 49(Pt 4):247-257 doi:10.1042/BA20070032

Tigges M, Fussenegger M (2006) Xbp1-based engineering of secretory capacity enhances the productivity of Chinese hamster ovary cells. Metab Eng 8(3):264-272 doi:10.1016/j.ymben.2006.01.006

Tjio JH, Puck TT (1958) Genetics of somatic mammalian cells. II. Chromosomal constitution of cells in tissue culture. J Exp Med 108(2):259-268

Tonomura A, Yerganian G (1956) Aneuploidy in the regenerating liver of the Chinese hamster. Genetics 41(5):664-665

Trummer E, Ernst W, Hesse F, Schriebl K, Lattenmayer C, Kunert R, Vorauer-Uhl K, Katinger H, Muller D (2008) Transcriptional profiling of phenotypically different Epo-Fc expressing CHO clones by cross-species microarray analysis. Biotechnol J 3(7):924-937 doi:10.1002/biot.200800038

Urlaub G, Chasin LA (1980) Isolation of Chinese hamster cell mutants deficient in dihydrofolate reductase activity. Proc Natl Acad Sci U S A 77(7):4216-4220 doi:10.1073/pnas.77.7.4216

Urlaub G, Käs E, Carothers AM, Chasin LA (1983) Deletion of the diploid dihydrofolate reductase locus from cultured mammalian cells. Cell 33(2):405-412 doi:10.1016/0092-8674(83)90422-1

Vishwanathan N, Le H, Jacob NM, Tsao YS, Ng SW, Loo B, Liu Z, Kantardjieff A, Hu WS (2014) Transcriptome dynamics of transgene amplification in Chinese hamster ovary cells. Biotechnol Bioeng 111(3):518-528 doi:10.1002/bit.25117

Vlachos IS, Kostoulas N, Vergoulis T, Georgakilas G, Reczko M, Maragkakis M, Paraskevopoulou MD, Prionidis K, Dalamagas T, Hatzigeorgiou AG (2012) DIANA miRPath v.2.0: investigating the combinatorial effect of microRNAs in pathways. Nucleic Acids Res 40(Web Server issue):W498-W504 doi:10.1093/nar/gks494

Volmer M, Northoff S, Scholz S, Thute T, Buntemeyer H, Noll T (2011) Fast filtration for metabolome sampling of suspended animal cells. Biotechnol Lett 33(3):495-502 doi:10.1007/s10529-010-0466-7

Waegeman H, Soetaert W (2011) Increasing recombinant protein production in *Escherichia coli* through metabolic and genetic engineering. J Ind Microbiol Biotechnol 38(12):1891-1910 doi:10.1007/s10295-011-1034-4

Walsh G (2010) Biopharmaceutical benchmarks 2010. Nat Biotechnol 28(9):917-924 doi:10.1038/nbt0910-917

Walter P, Ron D (2011) The unfolded protein response: from stress pathway to homeostatic regulation. Science 334(6059):1081-1086 doi:10.1126/science.1209038

Wang J, Huang W, Wu Y, Hou J, Nie Y, Gu H, Li J, Hu S, Zhang H (2012) MicroRNA-193 pro-proliferation effects for bone mesenchymal stem cells after low-level laser irradiation treatment through inhibitor of growth family, member 5. Stem Cells Dev 21(13):2508-2519 doi:10.1089/scd.2011.0695

Westwood AD, Rowe DA, Clarke HR (2010) Improved recombinant protein yield using a codon deoptimized DHFR selectable marker in a CHEF1 expression plasmid. Biotechnol Prog 26(6):1558-1566 doi:10.1002/btpr.491

Wiebe ME, Becker F, Lazar R, May L, Casto B, Semense M, Fautz C, Garnick R, Miller C, Masover G, Bergmann D, Lubiniecki AS (1989) A multifaceted approach to assure that recombinant tPA is free of adventitious virus. In: Spier RE, Griffiths JB, Stephenne J, Crooy PJ (eds) Advances in Animal Cell Biology and Technology for Bioprocesses. Butterworth-Heinemann, London, pp 68-71

Wlaschin KF, Hu WS (2007) Engineering cell metabolism for high-density cell culture via manipulation of sugar transport. J Biotechnol 131(2):168-176 doi:10.1016/j.jbiotec.2007.06.006

Wlaschin KF, Nissom PM, Gatti Mde L, Ong PF, Arleen S, Tan KS, Rink A, Cham B, Wong K, Yap M, Hu WS (2005) EST sequencing for gene discovery in Chinese hamster ovary cells. Biotechnol Bioeng 91(5):592-606 doi:10.1002/bit.20511

Wong DC, Wong KT, Lee YY, Morin PN, Heng CK, Yap MG (2006) Transcriptional profiling of apoptotic pathways in batch and fed-batch CHO cell cultures. Biotechnol Bioeng 94(2):373-382 doi:10.1002/bit.20872

Wu M, Shen Q, Yang Y, Zhang S, Qu W, Chen J, Sun H, Chen S (2013) Disruption of *YPS1* and *PEP4* genes reduces proteolytic degradation of secreted HSA/PTH in *Pichia pastoris* GS115. J Ind Microbiol Biotechnol 40(6):589-599 doi:10.1007/s10295-013-1264-8

Wu TY, Zhang TH, Qu LM, Feng JP, Tian LL, Zhang BH, Li DD, Sun YN, Liu M (2014) MiR-19a is correlated with prognosis and apoptosis of laryngeal squamous cell carcinoma by regulating TIMP-2 expression. Int J Clin Exp Pathol 7(1):56-63

Wurm FM (2013) CHO Quasispecies—Implications for Manufacturing Processes. Processes 1(3):296-311 doi:10.3390/pr1030296

Wurm FM, Hacker D (2011) First CHO genome. Nat Biotechnol 29(8):718-720 doi:10.1038/nbt.1943

Xing ZZ, Li ZJ, Chow V, Lee SS (2008) Identifying inhibitory threshold values of repressing metabolites in CHO cell culture using multivariate analysis methods. Biotechnology Progress 24(3):675-683 doi:10.1021/Bp070466m

Xu N, Brodin P, Wei T, Meisgen F, Eidsmo L, Nagy N, Kemeny L, Stahle M, Sonkoly E, Pivarcsi A (2011a) MiR-125b, a microRNA downregulated in psoriasis, modulates keratinocyte proliferation by targeting FGFR2. J Invest Dermatol 131(7):1521-1529 doi:10.1038/jid.2011.55

Xu X, Nagarajan H, Lewis NE, Pan S, Cai Z, Liu X, Chen W, Xie M, Wang W, Hammond S, Andersen MR, Neff N, Passarelli B, Koh W, Fan HC, Wang J, Gui Y, Lee KH, Betenbaugh MJ, Quake SR, Famili I, Palsson BO, Wang J (2011b) The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. Nat Biotechnol 29(8):735-741 doi:10.1038/nbt.1932

Yamaguchi A, Hori O, Stern DM, Hartmann E, Ogawa S, Tohyama M (1999) Stress-associated endoplasmic reticulum protein 1 (SERP1)/Ribosome-associated membrane protein 4 (RAMP4) stabilizes membrane proteins during stress and facilitates subsequent glycosylation. J Cell Biol 147(6):1195-1204 doi:10.1083/jcb.147.6.1195

Yamamoto K, Sato T, Matsui T, Sato M, Okada T, Yoshida H, Harada A, Mori K (2007) Transcriptional induction of mammalian ER quality control proteins is mediated by single or combined action of ATF6alpha and XBP1. Dev Cell 13(3):365-376 doi:10.1016/j.devcel.2007.07.018

Ye J, Alvin K, Latif H, Hsu A, Parikh V, Whitmer T, Tellers M, de la Cruz Edmonds MC, Ly J, Salmon P, Markusen JF (2010) Rapid protein production using CHO stable transfection pools. Biotechnol Prog 26(5):1431-1437 doi:10.1002/btpr.469

Ye JX, Ly J, Watts K, Hsu A, Walker A, McLaughlin K, Berdichevsky M, Prinz B, Kersey DS, d'Anjou M, Pollard D, Potgieter T (2011) Optimization of a glycoengineered *Pichia pastoris* cultivation process for commercial antibody production. Biotechnology Progress 27(6):1744-1750 doi:10.1002/Btpr.695

Yee JC, de Leon Gatti M, Philp RJ, Yap M, Hu WS (2008a) Genomic and proteomic exploration of CHO and hybridoma cells under sodium butyrate treatment. Biotechnol Bioeng 99(5):1186-1204 doi:10.1002/bit.21665

Yee JC, Gerdtzen ZP, Hu WS (2009) Comparative transcriptome analysis to unveil genes affecting recombinant protein productivity in mammalian cells. Biotechnol Bioeng 102(1):246-263 doi:10.1002/bit.22039

Yee JC, Wlaschin KF, Chuah SH, Nissom PM, Hu WS (2008b) Quality assessment of cross-species hybridization of CHO transcriptome on a mouse DNA oligo microarray. Biotechnol Bioeng 101(6):1359-1365 doi:10.1002/bit.21984

Yerganian G (1952) Cytogenetic possibilities with the Chinese hamster, Cricetulus barabensis griseus. Genetics 37(5):638-639

Yerganian G (1985) The biology and genetics of Chinese hamster. In: Gottesmann M (ed) Molecular Cell Cenetics. Wiley, New York, pp 3-36

Yoshikawa T, Nakanishi F, Ogura Y, Oi D, Omasa T, Katakura Y, Kishimoto M, Suga KI (2001) Flow cytometry: an improved method for the selection of highly productive gene-amplified CHO cells using flow cytometry. Biotechnol Bioeng 74(5):435-442 doi:10.1002/bit.1134

Young TM, Tsai M, Tian B, Mathews MB, Pe'ery T (2007) Cellular mRNA activates transcription elongation by displacing 7SK RNA. PLOS ONE 2(10):e1010 doi:10.1371/journal.pone.0001010

Zahn-Zabal M, Kobr M, Girod PA, Imhof M, Chatellard P, de Jesus M, Wurm F, Mermod N (2001) Development of stable cell lines for production or regulated expression using matrix attachment regions. J Biotechnol 87(1):29-42 doi:10.1016/S0168-1656(00)00423-5

Zauner G, Selman MHJ, Bondt A, Rombouts Y, Blank D, Deelder AM, Wuhrer M (2013) Glycoproteomic analysis of antibodies. Molecular & Cellular Proteomics 12(4):856-865 doi:10.1074/mcp.R112.026005

Zhang F, Sun X, Yi X, Zhang Y (2006) Metabolic characteristics of recombinant Chinese hamster ovary cells expressing glutamine synthetase in presence and absence of glutamine. Cytotechnology 51(1):21-28 doi:10.1007/s10616-006-9010-y

Zhang GF, Sadhukhan S, Tochtrop GP, Brunengraber H (2011) Metabolomics, pathway regulation, and pathway discovery. J Biol Chem 286(27):23631-23635 doi:10.1074/jbc.R110.171405

Zhang L, Li X, Ke Z, Huang L, Liang Y, Wu J, Zhang X, Chen Y, Zhang H, Luo X (2013) MiR-99a may serve as a potential oncogene in pediatric myeloid leukemia. Cancer Cell Int 13(1):110 doi:10.1186/1475-2867-13-110

Zheng YS, Zhang H, Zhang XJ, Feng DD, Luo XQ, Zeng CW, Lin KY, Zhou H, Qu LH, Zhang P, Chen YQ (2012) MiR-100 regulates cell differentiation and survival by targeting RBSP3, a phosphatase-like tumor suppressor in acute myeloid leukemia. Oncogene 31(1):80-92 doi:10.1038/onc.2011.208

Zhu H, Shyh-Chang N, Segre AV, Shinoda G, Shah SP, Einhorn WS, Takeuchi A, Engreitz JM, Hagan JP, Kharas MG, Urbach A, Thornton JE, Triboulet R, Gregory RI, Consortium D, Investigators M, Altshuler D, Daley GQ (2011) The *Lin28/let-7* axis regulates glucose metabolism. Cell 147(1):81-94 doi:10.1016/j.cell.2011.08.033

Appendix

Supporting information: Validation of CHO-specific microarray

Supporting material and methods

Gene expression determination by qRT-PCR

Transcript levels were measured by one-step quantitative reverse transcription PCR (qRT-PCR) using SYBR Green I. At first, total RNA extracts were treated with RQ1 RNase-free DNase (Promega, WI, USA) according to the manufacturer's instructions in order to remove potential residues of genomic DNA. qRT-PCR was performed on a Chromo4 real-time PCR system (Bio-Rad, CA, USA) using the QuantiFast SYBR Green RT-PCR Kit (Qiagen, Germany) according to the supplier's manual. Expression levels were normalized using two internal references (*Gapdh* and *Actr5*) to improve the reliability of the assay. *Gapdh* was highly and *Actr5* moderately expressed and the transcript levels for both genes were very stable in the microarray experiment across all CHO cell lines used in this study. Primers were preferentially designed to span an exon-exon junction or on two different exons, leading to an amplicon size between 100 and 200 bp. All primers used for this experiment are listed in Table 3. A 20 µL reaction mix contained 10 ng total RNA, 1 µM of each oligonucleotide primer, 0.2 µL QuantiFast RT Mix and 10 µL 2× QuantiFast SYBR Green RT-PCR Master Mix. qRT-PCR was performed at 50°C for 10 min (reverse transcription) and 95°C for 5 min (initial polymerase activation) followed by 40 cycles of 95°C for 10 s and 60°C for 30 s. Fluorescence was recorded after each amplification cycle. The specificity of the reaction was verified by analyzing the melting curve after the last cycle. Quantification cycle (C_q) values were determined with the software Opticon Monitor 3.1.32 (Bio-Rad). Relative expression ratios were computed using the software REST 2009, which allowed the consideration of both internal references and uses a complex Taylor series to estimate the standard error (SE).

Supporting Tables and Figures

Table 3: Primers used to determine transcription levels by qRT-PCR

Primer	Sequence (5'–3')
Acsl1_F	TGAACCGGATGTTTGACAGA
Acsl1_R	ATCCCACAGGCTGTTGTTTC
Actr5_F	CTCCTTCCAGGTTCAGCTTG
Actr5_R	GGCACAATGTTCCTTGAGGT
Adh5_F	GCCCATCCAGGAAGTACTCA
Adh5_R	GATTTCTTCACCCGAAGCAG
Akr1b7_F	CGGATACCGCCACATAGACT
Akr1b7_R	CCAGTTTCAGATCCGAGAGG
Arhgap24_F	CCTCCTTTGACAGCAACACA
Arhgap24_R	TTCCTTCACACCTGCTTCCT
Ccl2_F	ACGCTTCTGTGCCTGCTACT
Ccl2_R	TCCTCTTGGCGGTAAATGAG
Cd68_F	CAGCTCCAAGCACAAATTCA
Cd68_R	CTGCTCCATGTCCTGTTTGA
Ces1f_F	CAACACTGTGCCCTACATCG
Ces1f_R	CAACTGGAACTGCATCCTCA
Cpt1a_F	GCAAAGGGCTGATCAAGAAG
Cpt1a_R	GAAAGAGCCGAGTCATGGAG
Crot_F	GAAACTGGCTGGAAGAGTGG
Crot_R	AACTGAGTGCCTTCCATTGC
Ctla2a_F	GGAATATGAGCAGGGCAAGA
Ctla2a_R	GGAATCTCTACCAATCCACAGC
Cxcl3_F	GTGCCTGAAGACCCTACCAA
Cxcl3_R	TGGACTTGGCCTTCTTCAGT
Dusp14_F	AGGCATTGCTCAGATCACCT
Dusp14_R	CGATGGTAGCATTCACGATG
Ereg_F	GTGCATCTACCTGGTGGACA
Ereg_R	TGGATCCAGCGATTATGACA
Fabp4_F	CCGATCAGAGAGCACATTCA
Fabp4_R	CTTGTCACCCTCTCGCTTTC
Gapdh_F	GTAAGAAGCCCACCCTGGA
Gapdh_R	GTGAGGGAGATGATCGGTGT
Glud1_F	GGTGAATCTGATGGGAGCAT
Glud1_R	GACTTTGGGTGCATTGGACT
Gnb1_F	CAAAGCTGACAGAGCAGGTG
Gnb1_R	TCAGGGTTGGAATGGTCTTC
Gstm1_F	ATCCCTTCCCAAACCTGAAG

Primer	Sequence (5'–3')
Gstm1_R	CTCCATATGAGAGCCCGTGT
Il15_F	AGGCTGGCATTCATGTCTTC
Il15_R	TGCAACTGGGATGAAAGTCA
Lipa_F	ACCTCGGTGCAAAACTTACG
Lipa_R	CAAAGCAGTGGGTACAAGCA
Lrp1_F	ACCACCAGCTACCTCATTGG
Lrp1_R	CCTGGCCACACTAATGGTCT
Mmp12_F	TCCTGTCATTCCATCAGCAG
Mmp12_R	ATCCACCAGACGAACCTGTC
Nadk_F	CACTGCATGGGTGTCTTTTG
Nadk_R	GCAAGGCTCTCAAACCAGTC
Ppap2a_F	CCTTCCAACGAGGAGTGTTC
Ppap2a_R	GGTGCCAAAGTGAACAGACA
Rpl15_F	TGTTGCTGAGGAGAGAGCTG
Rpl15_R	TCTCTGTGCTTGTGGACTGG
Sbsn_F	ATGCTCATAACGGGGTCAAC
Sbsn_R	GCGATACTCCTCCACAGAGC
Sgk1_F	ATCGCCAACAACTCCTATGC
Sgk1_R	TGCCCTTTCCAATCACTTTC
Slpi_F	GGTGCTGCCAATTTAGGTGT
Slpi_R	CCTATTGGGAGGATTCAGCA
Thoc7_F	CTCTGGGCAAAGAATTGGAG
Thoc7_R	GAGCTTCGTCCACCTCTGAC
Ugt1a1_F	TGTCCTGGAAATGACTGCTG
Ugt1a1_R	CAGGTCCAGAGGCTCTATCG
Upb1_F	TGGTCTCTCACGCAATCAAG
Upb1_R	ACATCTCAAGTCGGCCTGTC
Wapal_F	CTCCTACACAGCCCTGCTTC
Wapal_R	GACTTCTGGCCTGTTGTTCC

List of figures

Figure 1. US sales (billion US$) of the top nine categories of biological drugs in 2012 8
Figure 2. From Chinese hamster to Chinese hamster ovary (CHO) suspension cells. 12
Figure 3. Typical process to develop a recombinant high producer CHO cell line by gene amplification 15
Figure 4. MicroRNA biogenesis and their biotechnological application 21
Figure 5. Overview of omics techniques. 24
Figure 6. Schematic structure of a single-chain Fv-Fc fusion antibody 89
Figure 7. Total protein content in the supernatant during fed batch cultivation. 91
Figure 8. Principle of steady-state cultivation using a chemostat 94
Figure 9. Comparison of gene expression determined by qRT-PCR and microarray (*Acsl1–Cd68*) 96
Figure 10. Comparison of gene expression determined by qRT-PCR and microarray (*Cesf1f–Dusp14*) 97
Figure 11. Comparison of gene expression determined by qRT-PCR and microarray (*Ereg–Il15*) 98
Figure 12. Comparison of gene expression determined by qRT-PCR and microarray (*Lipa–Rpl15*) 99
Figure 13. Comparison of gene expression determined by qRT-PCR and microarray (*Sbsn–Upb1*) 100
Figure 14. Comparison of gene expression determined by qRT-PCR and microarray (*Wapal*) 101
Figure 15. Number of commonly and exclusively differentially expressed genes 103
Figure 16. Principal component analysis (PCA) of the whole dataset (8 comparisons) 104
Figure 17. The three branches of unfolded protein response (UPR) in metazoans 106
Figure 18. Expression profile of potential $ATF6_N$ target genes (ATF6 branch of UPR) 107
Figure 19. Expression profile of potential ATF4 target genes (PERK branch of UPR) 108
Figure 20. Expression profile of potential XBP1s target genes (IRE1 branch of UPR) 109
Figure 21. Correlation of transgene copy number, product mRNA level and specific product secretion rate (q_P) 110
Figure 22. Expression profile of transcription factors associated with the cytomegalovirus (CMV) major immediate-early (MIE) enhancer/promoter 111
Figure 23. Number of commonly and exclusively differentially expressed miRNAs 115

Abbreviations

μ	specific growth rate
2D-DIGE	Two-dimensional difference gel electrophoresis
2D-PAGE	two-dimensional polyacrylamide gel electrophoresis
3D6_H	CHO 3D6scFv-Fc high producer
3D6_L	CHO 3D6scFv-Fc low producer
3D6scFv-Fc	3D6 anti-HIV-1 single-chain Fv-Fc fusion antibody
ADCC	antibody-dependent cellular cytotoxicity
adj. *p*	adjusted *p*-value
AGO	Argonaute
ARE	AU-rich element
ATF6	activating transcription factor 6
BAC	bacterial artificial chromosomes
CDC	complement-dependent cytotoxicity
CDK2	cyclin-dependent kinase 2
CDKL3	cyclin-dependent kinase like 3
CDM	cell dry mass
CHEF-1	CHO-derived elongation factor-1α
CHO	Chinese hamster ovary
CIRP	cold-inducible RNA-binding protein
CMV	cytomegalovirus
CNV	copy number variation
CPS I	carbamoyl phosphate synthetase I
C_q	quantification cycle
CTD	carboxy-terminal domain
CTDSPL	CTD small phosphatase-like protein
D	dilution rate
DHFR	dihydrofolate reductase
DMSO	dimethyl sulfoxide
E. coli	*Escherichia coli*
ER	endoplasmic reticulum
ERAD	ER-associated degradation
EST	expressed sequence tag
EV	CHO empty vector (non-producer)
F	flow rate
FACS	fluorescence-activated cell sorting
FC	fold change
Fc	fragment crystallizable
GO	Gene Ontology
GOI	gene of interest
GPDH	glycerol-3-phosphate dehydrogenase
GS	glutamine synthetase
HCDC	high cell density cultivation

HCP	host cell protein
HIC	human I-mfa domain containing protein
HIV	human immunodeficiency virus
HMG-CoA-S	3-hydroxy-3-methylglutaryl-CoA synthase
HPLC	high-performance liquid chromatography
HSA	human serum albumin
HSA_H	CHO HSA high producer
HSA_L	CHO HSA low producer
IL-15	interleukin-15
indel	insertion and deletion
ING5	inhibitor of growth family 5
IRE1	inositol requiring enzyme 1
IRES	internal ribosome entry site
iTRAQ	isobaric tags for relative and absolute quantification
IVCC	integral of viable cell concentration
JNK	c-Jun N-terminal kinase
LDH	lactate dehydrogenase
LPS	lipopolysaccharides
LTR	long terminal repeat
mAb	monoclonal antibody
MAPK	mitogen activated protein kinase
MAR	matrix attachment region
MDH II	malate dehydrogenase II
MIE	major immediate-early
miRISC	miRNA-induced silencing complex
miRNA	microRNA
MODC	mouse ornithine decarboxylase
MS	mass spectrometry
MS/MS	tandem mass spectrometry
MSC	mesenchymal stem cell
MSX	methionine sulphoximine
mTOR	mammalian target of rapamycin
MTX	methotrexate
NF- κB	nuclear factor-kappa B
NFIB	nuclear factor I/B
NGS	generation sequencing
NMR	nuclear magnetic resonance
OTC	ornithine transcarbamoylase
P. pastoris	*Pichia pastoris*
$p27^{Kip1}$	cyclin-dependent kinase inhibitor 1B
$p57^{Kip2}$	cyclin-dependent kinase inhibitor 1C
PCA	principal component analysis
PDI	protein disulfide isomerase
PERK	double-stranded RNA-activated protein kinase-like ER kinase

piRNAs	PIWI-interacting RNAs
P-TEFb	positive transcription elongation factor b
PTM	post-translational modification
q_P	specific product secretion rate
qRT-PCR	quantitative reverse transcription polymerase chain reaction
RAMP4	ribosome-associate membrane protein 4
rDNA	recombinant DNA
RISC	RNA-induced silencing complex
RMCE	recombinase-mediated cassette exchange
RNA Pol II	RNA Polymerase II
RNA-Seq	transcriptome (RNA) sequencing
RSV	Rous sarcoma virus
S	substrate concentration
S. cerevisiae	*Saccharomyces cerevisiae*
SAR	scaffold attachment region
SD	standard deviation
SE	standard error
SERP1	stress-associated endoplasmic reticulum protein 1
SILAC	stable isotope labeling
SNARE	Soluble NSF (N-ethylmaleimide-sensitive factor) receptor
SNP	single-nucleotide polymorphism
SRP	signal recognition particle
STY	space-time yield
SV40	Simian virus 40
TCA	tricarboxylic acid
TIMP-2	Tissue inhibitor of metalloproteinase 2
TMT	tandem mass tag
TPO	thrombopoietin
TRIB2	tribbles pseudokinase 2
UCOE	Ubiquitous chromatin opening element
UPR	unfolded protein response
UTR	untranslated region
V	culture volume
VCC	viable cell concentration
VCP	valosin-containing protein
VH	variable heavy chain
VL	variable light chain
x	cell concentration
XBP1	X-box binding protein 1

Curriculum vitae

Personal details

Name:	Andreas Michael Maccani
Academic degree:	Dipl.-Ing., Bakk.techn.
Current address:	Montleartstraße 70/22, 1160 Vienna, Austria
Phone:	+436508623740
E-mail:	andreas.maccani@gmx.net
Date of birth:	22 June 1977
Place of birth:	Bregenz, Austria
Nationality:	Austria

Education:

2010 – 2014: PhD program "Biomolecular Technology of Proteins (BioToP)" at the University of Natural Resources and Life Sciences, Vienna, Austria

2007 – 2010: Master in Biotechnology at the University of Natural Resources and Life Sciences, Vienna, Austria
Passed with distinction on 26 February 2010

2004 – 2007: Bachelor in Food Science and Biotechnology at the University of Natural Resources and Life Sciences, Vienna, Austria
Passed with distinction on 12 October 2007

1991 – 1996: Technical school for electronics, HTL Rankweil, Austria

1987 – 1991: Grammar school

1983 – 1987: Primary school

Employment history:

April 2010 – present: Junior Researcher at the Austrian Centre of Industrial Biotechnology (ACIB), University of Natural Resources and Life Sciences, Vienna, Austria

October 2006 – January 2010: Tutor at the Institute of Applied Statistics and Computing at the University of Natural Resources and Life Sciences, Vienna, Austria

June 2008 – September 2008: Internship at Boehringer Ingelheim, Vienna, Austria

August 2007: Internship at the Österreichische Agentur für Gesundheit und Ernährungssicherheit (Austrian Agency for Health and Food Safety), Vienna, Austria

August 2005 – September 2005: Internship at the Umweltinstitut des Landes Vorarlberg (Environmental Institute of the State government of Vorarlberg), Bregenz, Austria

July 2002 – September 2004: Electronic engineer at Tridonic, Dornbirn, Austria

September 2000 – April 2002: Electronic engineer at Newtronics, Melbourne, Australia

June 1997 – August 2000: Electronic engineer at Tridonic, Dornbirn, Austria

July 1994 – August 1994: Internship at the Vorarlberger Kraftwerke, Bregenz, Austria

July 1993 – August 1993: Internship at the Elektrostudio Kreil & Lecher, Dornbirn, Austria

Research stays:

June 2013 – September 2013: Project: *qRT-PCR validation of microarray data*, Group of Professor Mark Smales, University of Kent, Canterbury, UK

Language skills:

German (native language)
English (fluent)
Italian (basic knowledge)

Publications:

Maccani A., Hackl M., Leitner C., Steinfellner W., Graf A.B., Tatto N.E., Karbiener M., Scheideler M., Grillari J., Mattanovich D., Kunert R., Borth N., Grabherr R., Ernst W. (2014) Identification of microNRAs specific for high producer CHO cell lines using steady-state cultivation. *Applied Microbiology and Biotechnology*, in press

Maccani A., Landes N., Stadlmayr G., Maresch D., Leitner C., Maurer M., Gasser B., Ernst W., Kunert R., Mattanovich D. (2014) Pichia pastoris secretes recombinant proteins less efficiently than Chinese hamster ovary cells but allows higher space-time yields for less complex proteins. *Biotechnology Journal* 9, 526-537

Maccani A., Ernst W., Grabherr R. (2013) Whole genome sequencing improves estimation of nuclear DNA content of Chinese hamster ovary cells. *Cytometry A* 83, 893-895

Krammer F., Pontiller J., Tauer C., Palmberger D., **Maccani A.**, Baumann M., Grabherr R. (2010) Evaluation of the influenza A replicon for transient expression of recombinant proteins in mammalian cells. *PLOS ONE* 5 (10), e13265

Pontiller J., **Maccani A**., Baumann M., Klancnik I. and Ernst W. (2010) Identification of CHO endogenous gene regulatory elements. *Molecular Biotechnology* 45 (3), 235-240

Maccani A. (2010) Identification and characterisation of CHO endogenous gene regulatory elements. Master thesis, University of Natural Resources and Life Sciences, Vienna

Poster presentations at conferences:

Comparative microRNA expression profiling of recombinant Chinese hamster ovary cell lines and the challenge of target identification [CHOgenome Workshop 2014 "Genome Scale Science for CHO Production Cells: Genomes, -Omics and Big Data", Vienna, Austria, March 13 – 14, 2014]

Messenger RNA and microRNA expression profiling of high and low producing recombinant Chinese hamster ovary cell lines [23rd ESACT Meeting, Lille, France, June 22 – 26, 2013]

Transcriptome study of high and low producing recombinant Chinese hamster ovary cell lines [7th Conference on Recombinant Protein Production – RPP7, Laupheim, Germany, March 6 – 8, 2013]

Comparative transcriptome analysis of recombinant high and low producing Chinese hamster ovary cell lines using steady-state cultivation [4nd Annual ÖGMBT Meeting, Graz, Austria, 17 – 19 September, 2012]

Steady-state cultivation of Chinese hamster ovary cells for comparative physiological analyses [Cell Culture Engineering XIII, Scottsdale, USA, 23 – 27 April, 2012]