



## MASTER THESIS

Title

Analysis of Expression and Inhibition of microRNAs in CHO Cell Factories

Author: Matthias Hackl  
Matrikel-Number: 0340208  
Field of Study: 066 418 Biotechnologie  
Supervisors: Priv. Doz. Dr. Dethardt Müller  
Priv. Doz. DI Dr. Johannes Grillari

Vienna, March 2009

*meinen lieben Eltern*

## *Acknowledgement*

The last 5  $\frac{1}{2}$  years of studying at BOKU University have been a wonderful learning experience for me due to several facts: above all the continuous support through my parents and especially their way of encouraging me to pursue my interests and to realize my plans and dreams (including those that did not concern my studies) made my years of studying an enjoyable journey. There is no way I could have done this without them. Second the great guidance of the faculty members and professors I met and got to know during my studies at BOKU, especially Iain Wilson during my Bachelor years, and Dethardt Müller and Johannes Grillari during my Master's degree, raised my interest for many scientific topics. Their great motivation made it easy for me to work hard and learn a lot and the constant and immediate help of Dethardt and Johannes during my research was vital for driving my understanding and interest in my topic. I also want to thank the Austrian Center of Biopharmaceutical Technology (ACBT) for the kind support, all members of the Mammalian Cell Culture group at BOKU, especially Lukas Fiedler, Georg Hinterkörner, Johannes Pichler and Thomas Sterovsky for their patience and help. Also several members of the Aging Research group, especially Matthias Wieser, Carina Schreiner, Johanna Bisich and Harald Kühnel greatly supported me throughout my work and involved me in healthy discussions on microRNAs.

I have to thank Max Goritschnig and Professor Hermann Katinger for their help in organizing my stay at the University of Minnesota and the Marshallplan Stiftung for funding me. During the 3 months I spent at Prof. Wei-Shou Hu's lab, it was his and his graduate student's kind support that made me feel very welcome and allowed me to productively work on my research project. Especially working together with Nitya Jacob, Kathryn Johnson, Bhanu Mulukutla and Anne Kantjardeff made my stay in Minnesota a wonderful experience, from a scientific as well as personal perspective. In this respect I also have to thank Eugene Allred for hosting me in Minneapolis and for his support and friendliness.

Finally I want to thank all my friends from Club Biotech, especially Friedrich Erhart and Martin Kleinheinz for the healthy discussion and the philosophical support, and above all my better half, Andrea Temper, for her great patience, constant support and true love.

## *Abstract*

In a time where expiring intellectual property rights for recombinant protein therapeutics have enhanced competition among Biotech companies, increasing the efficiency of bioprocesses has become more important than ever. Manifold approaches to increase space-time yield of a mammalian cell based process are available. Most commonly they are based on improving the external environment of cells due to highly sophisticated process design, or on changing the intracellular setup of cells, for example via genetic engineering. Recently the latter approach has seen an extension of possibilities due to the discovery of powerful RNAi regulators, referred to as microRNAs.

MicroRNAs are a class of small noncoding RNA molecules that exhibit post-transcriptional regulation of messenger RNAs. In most cases interactions of miRNAs with their target mRNAs result in negative regulation, in vertebrates mainly due to translational repression. For almost 20 years now, the impact of microRNAs on several cellular pathways has been studied, and with over 500 miRNA species identified in humans so far, there is no major pathway left which is not affected by miRNAs. Special emphasis has been put on discovering disease related roles of miRNAs, resulting in the identification of a set of *oncomirs* which are commonly overexpressed in tumor cells and frequently regulate cell proliferation or cell death as well as cell transformation and invasion. Hence these regulators are of great potential use for altering biopharmaceutical work horses, such as Chinese Hamster Ovary, into fast growing, high producing and stress resistant cell factories.

The aim of this study was to enlarge the knowledge about microRNAs and their regulatory roles in Chinese Hamster Ovary cells. We applied deep sequencing technology to identify 260 distinct Chinese Hamster miRNAs solely based on sequence conservation in 25 species. We found the highest degree of conservation in human and rodent species and elaborated differences between several CHO cell lines in terms of their miRNA expression pattern. The results indicated that changes in miRNA expression are more likely to occur on account of genetic changes during cell line development such as gene amplification than due to addition of small molecule enhancers or changes in cultivation conditions.

Based on putative identification of PDCD4 as a target of miR-21 in Chinese Hamster, and the public available sequence for miR-21 in CHO (cgr-miR-21), expression and inhibition studies were performed to assess the impact of miR-21 on cell growth, productivity as well as stress resistance in a CHO production clone. It was found that under the tested conditions miR-21 did not cause significant changes in cell phenotype, maybe due to a different genetic setup of CHO cells compared to malignant cancer cell models.

In order to identify additional microRNA targets for CHO cell engineering, profiling techniques such as miRNA microarrays were applied. Therefore RNA from sodium butyrate and temperature shifted cultivations was extracted as well as from CHO cells undergoing methotrexate gene amplification, but samples remained at the University of Minnesota for further analysis. At BOKU University RNA from CHO temperature shift cultivations was used for miRNA profiling under low temperature conditions, which identified miR-21 as upregulated at 33°C. Eventually, information about miRNA expression was used to draw conclusions about the miRNA setup of CHO and to discuss perspectives of the use of miRNAs for CHO cell engineering.

## *Kurzfassung*

Angesichts des bevorstehenden Ablaufs von Patenten auf biotechnologisch hergestellte Wirkstoffe, steht ein zunehmend stärker werdender Konkurrenzkampf in der Biotechnologie bevor. Infolgedessen, um die Konkurrenzfähigkeit aufrecht zu erhalten, wird die Entwicklung von noch effektiveren Produktionsprozessen für viele Unternehmen noch weiter an Bedeutung gewinnen. Zwei Strategien zur Verbesserung der Raum-Zeit Ausbeute von *mammalian-cell based* Bioprozessen sind momentan etabliert: einerseits führt die Optimierung von Prozessführung und Mediumszusammensetzung zu besseren Wachstumsbedingungen von tierischen Zellen, andererseits können intrinsische Eigenschaften von Zellen, wie Wachstumskapazität und spezifische Produktivität verbessert werden, meist durch gentechnische Veränderung. Letzteres erfolgt zum Beispiel durch gezielte Veränderung von zellulären Programmen wie Zellteilung, Zelltod und Produktsekretion; Programme für die kürzlich bewiesen wurde dass sie unter anderem durch kleine RNA Spezies reguliert werden – sogenannten microRNAs.

Als circa 22 Nukleotide umfassende Moleküle gehören microRNAs zur Klasse der *kleinen nicht-kodierenden* RNAs, deren regulierende Eigenschaften auf einer meist negativen post-transkriptionellen Kontrolle der messenger RNA Translation beruhen. Mit über 500 bekannten microRNAs im Menschen, gibt es kaum einen zellulären Mechanismus der nicht unter dem Einfluss von microRNAs steht. Unter den regulierten zellulären Programmen finden sich viele die auch aus biotechnologischer Sicht interessant sind, zum Beispiel Zellteilung, Apoptose oder Proteinfaltung und Sekretion. Die Untersuchung von microRNAs, die als globale Regulatoren die Translation von bis zu 100 mRNAs kontrollieren, in Zellfabriken wie zum Beispiel Chinese Hamster Ovary (CHO) Zellen – ist naheliegend.

Das Ziel dieser Studie war es daher die regulierenden Eigenschaften von microRNAs in CHO zu untersuchen. Dafür wurden im ersten Schritt Sequenzieretechnologien der nächsten Generation angewendet, die Sequenzinformationen von 260 CHO microRNAs lieferten. Dabei wurde vor allem die starke Konservierung CHO miRNAs in nahen Verwandten wie Maus, Ratte aber auch Mensch deutlich. Mithilfe der Solexa-Daten war es auch möglich verschiedene Zelllinien

hinsichtlich ihrer microRNA Expression einzuordnen. Dabei wurde festgestellt dass vor allem genetische Änderungen in Zellen starke Unterschiede in der globalen microRNA expression hervorrufen, verglichen mit der Änderung von externen Bedingungen wie Temperatur oder der Zugabe von *small molecule enhancers*. Weiters gelang die Identifikation eines möglicherweise durch microRNA-21 regulierten Gens in CHO, programmed cell death 4 (PDCD4). Daher, und aufgrund der öffentlichen Verfügbarkeit der CHO Sequenz für miRNA-21, wurde mit Studien der Überexpression und Inhibierung dieser microRNA in CHO fortgeföhren. Jedoch konnten keine signifikanten Einflüsse von microRNA-21 auf Produktivität, Wachstum oder Stressresistenz festgestellt werden. Um weitere Kandidaten für Überexpressions- und Knockdownstudien zu finden, wurden microRNA Microarrays verwendet. Dafür wurden verschiedene biotechnologisch relevante Proben sowohl an der BOKU als auch an der University of Minnesota erzeugt, wie zum Beispiel Kälteschock behandelte Zellen oder mit Natriumbutyrat, ein Agens das die spezifische Produktivität der Zellen erhöht, behandelte Zellen. Die schon publizierte Hochregulierung von microRNA-21 unter niedrigen Temperaturen konnte dabei bestätigt werden. Weiters ermöglichten die gewonnen Daten ein besseres Verständnis von miRNAs in CHO und halfen bei der Entwicklung von neuen Strategien um microRNAs für die genetische Optimierung von CHO Produktionsklonen zu verwenden.

## Contents

<b>1. Introduction</b> .....	<b>1</b>
1.1. CHO - a Cell Factory for Modern Bioprocesses.....	1
1.2. Improving Industrial Bioprocesses.....	2
1.2.1. Medium Optimization.....	3
1.2.2. Bioprocess Optimization.....	4
1.2.3. Small Molecule Enhancers.....	5
1.2.4. DNA Delivery and Integration.....	5
1.2.5. Genetic Engineering of Host Cells.....	6
1.2.5.1. Apoptosis Engineering.....	7
1.3. Engineering Host Cells on microRNA Level.....	10
1.3.1. History of microRNAs.....	10
1.3.2. MicroRNA Genomics and Biogenesis.....	11
1.3.3. Function of miRNA as post-transcriptional Regulators.....	14
1.3.4. Relevance of microRNAs as Key Regulators in Animal Cells.....	15
1.3.5. MicroRNA-21.....	16
<b>2. Aims</b> .....	<b>18</b>
<b>3. Material and Methods</b> .....	<b>20</b>
3.1. Cell Culture.....	20
3.1.1. Cell Lines and Media Composition.....	20
3.1.2. Cell Culture Parameters.....	21
3.1.3. Temperature Shift and Sodium Butyrate Treatment.....	22
3.1.4. Methotrexate Gene Amplification.....	23
3.1.5. MicroRNA Precursor Transfection.....	24
3.2. Analytical Methods.....	26
3.2.1. Cell Number and Cell Viability.....	26
3.2.1.1. Hemocytometer.....	26
3.2.1.2. Automated Counting of Cell Number – Beckman Multisizer 3™.....	27
3.2.2. Flow Cytometry.....	28
3.2.2.1. Dual-Dye Measurements – Electronic Compensation.....	28
3.2.2.2. Viability – Propidium Iodide Stain.....	30

3.2.2.3.	Apoptosis – Annexin-V + Propidium Iodide Stain .....	31
3.2.2.4.	MicroRNA Transfection Efficiency .....	33
3.2.3.	Metabolite and Amino Acid Analysis .....	33
3.2.4.	EpoFc Product Titer Analysis - ELISA .....	33
3.3.	MicroRNA-21 Overexpression and Inhibition Assays.....	34
3.3.1.	Batch Cultivation .....	34
3.3.2.	Stress Resistance Assays .....	35
3.3.2.1.	Camptothecin Treatment.....	35
3.3.2.2.	Hydrogen Peroxide Treatment.....	36
3.3.2.3.	Nutrient Depletion Assay .....	37
3.4.	RNA Analysis .....	37
3.4.1.	Total RNA Isolation.....	37
3.4.2.	RNA Quality Assessment.....	38
3.4.3.	Total RNA Quantification .....	39
3.4.4.	Taqman quantitative real time PCR .....	40
3.5.	MicroRNA MicroArray Profiling .....	45
3.5.1.	Exiqon LNA MicroArrays (BOKU University) .....	45
3.5.1.1.	Array Design .....	45
3.5.1.2.	MicroRNA Labeling.....	46
3.5.1.3.	Hybridization .....	47
3.5.1.4.	Scanning .....	48
3.5.1.5.	Data Analysis.....	49
3.5.2.	Ambion mirVANA MicroArrays (University of Minnesota) .....	49
3.5.2.1.	Small RNA Isolation for Microarray Analysis.....	49
3.5.2.2.	MicroRNA labeling .....	50
3.5.2.3.	MicroRNA MicroArrays .....	50
3.5.2.4.	Slide Post Processing.....	50
3.5.2.5.	Hybridization .....	51
3.5.2.6.	Scanning .....	51
3.6.	Illumina/Solexa Sequencing .....	52
3.6.1.	Background .....	52
3.6.2.	Small RNA Library Preparation.....	54

3.6.3.	Small RNA Libraries .....	55
3.6.4.	Data Processing.....	56
3.6.5.	BLAST Identification of conserved miRNAs.....	56
<b>4.</b>	<b>Results .....</b>	<b>59</b>
4.1.	MicroRNA Sequencing .....	59
4.1.1.	The Solexa Dataset.....	59
4.1.2.	Annotation of Small RNAs .....	61
4.1.3.	Annotation of Conserved MicroRNAs .....	65
4.1.3.1.	Variability in microRNA processing - isomiRs .....	65
4.1.3.2.	Characteristics of microRNA Conservation in CHO .....	67
4.1.4.	Mature microRNA Expression Analysis .....	68
4.2.	MicroRNA Target Prediction in CHO Cells.....	71
4.3.	MicroRNA Overexpression and Inhibition in CHO .....	74
4.3.1.	Optimization of transient microRNA Transfection .....	74
4.3.2.	Real-Time qPCR Analysis of miR-21 .....	77
4.3.3.	Batch Cultivation of miR-21 Transfected CHO Cells.....	79
4.3.3.1.	Cell Growth .....	79
4.3.3.2.	Cell Productivity .....	81
4.3.4.	Stress Resistance of miR-21 Transfected Cells.....	81
4.3.4.1.	Camptothecin Treatment.....	81
4.3.4.2.	Nutrient Depletion .....	84
4.3.4.3.	Oxidative Stress – Hydrogen Peroxide Treatment.....	88
4.4.	MicroRNA Profiling Experiments .....	91
4.4.1.	Temperature Shift and Sodium Butyrate Treatment.....	91
4.4.2.	Methotrexate Gene Amplification .....	93
4.4.3.	Temperature Shift in four CHO Cell Lines – BOKU .....	94
4.4.3.1.	Real-Time qPCR Analysis .....	94
4.4.3.2.	Microarray Analysis – microRNA Expression Profiling.....	95
<b>5.</b>	<b>Discussion .....</b>	<b>100</b>
5.1.	MicroRNA Expression Profiling .....	100
5.2.	Functional Study of microRNA-21 in CHO.....	101

5.3. Perspectives .....	103
6. References .....	107
7. Appendix.....	114
7.1. Supplemental Table 1: Recombinant Protein Therapeutics produced in CHO.....	114

## 1. Introduction

### 1.1. CHO - a Cell Factory for Modern Bioprocesses

The use of mammalian derived cell lines for the production of therapeutic proteins has been increasing ever since the approval and release of human tissue plasminogen activator (tPA, Activase; Genentech, San Francisco, CA, USA) in 1987. Recombinant proteins have been proven to be of great importance in the treatment of a variety of human illnesses, ranging from cancer to infertility, and mammalian cell lines are the preferred cell factory used for their production (Wurm 2004). Today about 60% to 70% of recombinant proteins are expressed in mammalian cells, due to several advantages over other commonly used organisms such as the bacteria *E. coli* or *B. subtilis*, or yeast *P. pastoris*. In order to achieve effective treatment of patients, recombinant proteins must be synthesized in their biologically active forms, which means that certain quality criteria such as proper folding and post-translational modifications must be fulfilled. This includes specific addition of carbohydrates to certain amino-acid residues resulting in species-specific glycosylation patterns. Cell lines derived from human-tissues like HeLa or HEK293 or rodents, such as Chinese Hamster Ovary Cells (CHO), Baby Hamster Kidney (BHK) or NS0 (mouse myeloma) are naturally more suitable to fulfill these quality criteria than prokaryotic cells or single-cellular eukaryotes. However, also between these mammalian cell lines differences exist, which make certain cell lines preferred cell factories in Biotech industry. I.e. proteins expressed in CHO cells account for nearly 70% of all recombinant proteins produced resulting in annual sales of CHO derived products exceeding US\$ 30billion worldwide (Jayapal 2005). Supplementary Table 1 in the appendix, taken from the Consortium for Chinese Hamster Ovary Cell Genomics, views current CHO cell-derived therapeutic proteins.

One of the reasons why Chinese Hamster Ovary Cells turned from a mammalian cell model-organism used in fundamental research into the work-horse for modern recombinant protein production, was the isolation of several auxotrophs during mutagenesis studies on CHO. Based on deficiencies in metabolic enzymes such as APRT (adenine phosphoribosyl transferase) or DHFR (dihydrofolate reductase) mutants were derived which required the addition of several nutrients for maintaining growth and viability (Puck and Kao 1967; Taylor, Pipkorn et al. 1977;

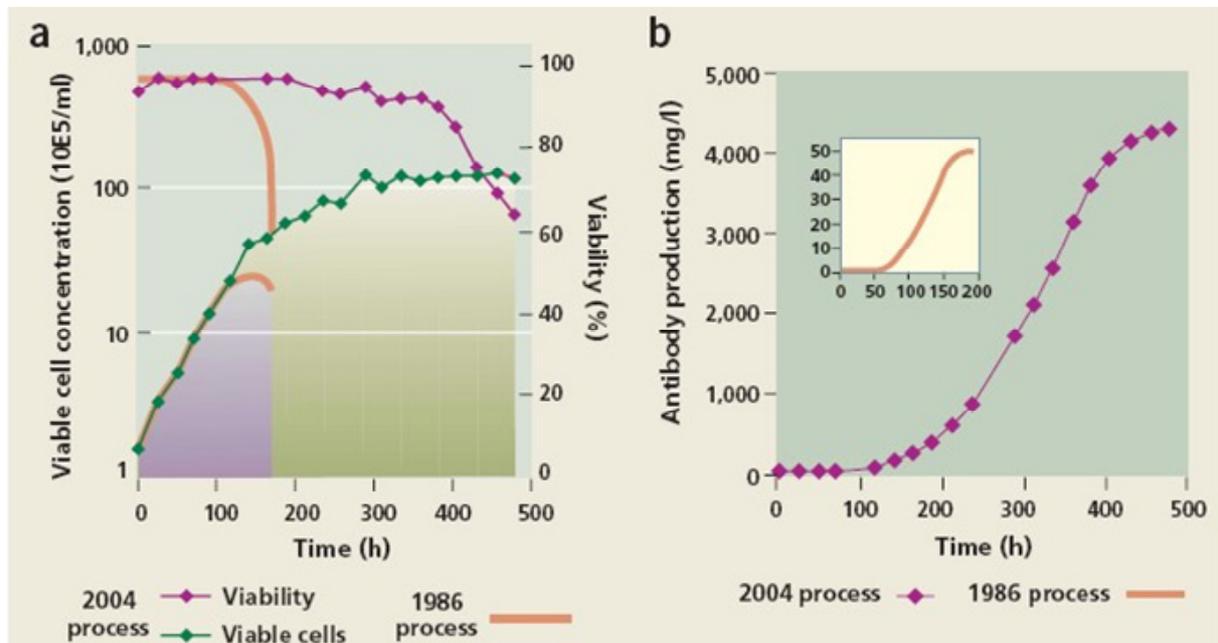
Urlaub and Chasin 1980). With the finding that this specific lack of a metabolic enzyme is of great use for transfecting, selecting, amplifying and stably expressing heterologous proteins, the use of these dhfr-negative parental cell lines (such as DG44 or DXB11) in biopharmaceutical industry saw a great boost (Pallavicini, DeTeresa et al. 1990; Gandor, Leist et al. 1995). As far as the issue of protein quality and efficacy is concerned, CHO derived recombinant proteins have been shown to carry glycosylation patterns that are compatible as well bioactive in humans. Also in respect to product safety - another big issue for biotech companies - CHO cells have advantages to human cell lines, as they are less likely to carry adventitious pathogenic agents that might find their way into human patients (Jayapal 2005). Finally the ability to adapt CHO cells to grow in suspension, as opposed to adherent growth, is highly appreciated since large stirred-tank bioreactors can be used for fermentation that can be scaled up more easily.

In spite of these advantages mentioned above, some concerns remained about the use of cultured mammalian cells for recombinant protein production. In explicit, the fact that altered expression of oncogenes caused these cell lines to transform, which means that they can escape senescence, leaves some doubts about product safety. However, these risks seem to be outweighed by the need for recombinant proteins as therapeutics and a 20 year track of safe use of CHO derived active pharmaceutical ingredients.

## **1.2. Improving Industrial Bioprocesses**

Despite numerous advantages of Chinese Hamster Ovary cell lines for the production of recombinant proteins, some major drawbacks exist that make this process time-consuming and very expensive. In general mammalian cell cultivations are characterized by slow growth rates and low specific productivity when compared to bacteria or yeast. Thus, overall low volumetric productivity (also referred to as space time yield or STY for short), which is a function of two basic culture parameters, cell specific production rate  $q_p$  and the integral of viable cell concentration, is low (Dinnis and James 2005; Trummer, Fauland et al. 2006). Also, demands on the composition of cell culture media are high, and the overall robustness of mammalian cells during a bioprocess is low. Consequently, over the last years considerable effort was put into developing strategies to address these problems. As reviewed by Florian Wurm, the combined

application of several strategies has led to a reasonable increase in product titers over the last 15 to 20 years (Wurm 2004).



**Figure 1 – Comparison of cell culture processes from 1986 and 2004:** the 1986 process is a hypothetical process based on unpublished results; the 2004 process corresponds to a 10L scale human recombinant antibody production process: a) viable cell concentration and viability is shown. Compared to 1986 recent bioprocess maintain high viabilities for up to 3 weeks, thus increasing the integral of viable cells (area under the VCC curve, referred to as cumulative cell days or CD). b) Antibody product titer is shown over the time-line of a bioprocess. Increase in culture duration and introduction of biphasic processes yields product titers up to 5 mg/L in 2004 compared to 50 µg/L in 1986. Figure taken from Wurm (2004).

### 1.2.1. Medium Optimization

One strategy is to provide cells with optimal surroundings for high productivity, for instance via improving media composition. While there are several excellent commercially available media for cell culture, leading manufacturers of recombinant proteins put considerable work in developing their own proprietary media formulations. Commonly manufacturers develop different media compositions for each specific growth phase within a bioprocess. For example formulations designed to support rapid growth within 3-5 days sub-culturing cycles are different from a production media that is used for batch or extended batch processes. Also media compositions have to be tuned in order to match the needs of each individual cell line.

### 1.2.2. Bioprocess Optimization

Biotech companies in cooperation with public research facilities also invested considerable work in the development of innovative fermentation processes, such as continuous fermentation based on cell-retention systems, which are referred to as perfusion processes. Briefly, as a continuous process, a perfusion system consists of a media influx and a suspension culture efflux. Within the efflux stream a cell retention system such as an ultrasonicator, hydro-cyclone or a hollow-fiber membrane module separates the cells from the culture broth, and recycles them into the bioreactor (Durr Schmid, Landauer et al. 2003). Thus, growth conditions within the reactor can be created that resemble a Batch process, while at the same time several reactor volumes of fresh media can be fed into the culture per day. This highly controlled process can be continued for several weeks up to months and high quality protein can be harvested in regular intervals. Alternatively extended batch cultures are used for high yield production processes, which are characterized by semi-continuous or batch-wise addition of media components (Wurm 2004).

Other developments in improving upstream bioprocesses introduced the concept of biphasic cultivation of mammalian cells. A biphasic bioprocess is characterized by an initial phase of fast growth resulting in accumulation of biomass, and a stationary phase characterized by high cell-densities, slow growth and high specific protein production. Thus, biphasic processes are based on the inverse relationship between growth rate  $\mu$  and specific production rate  $q_p$  (Fussenegger et al. 1998). This can be explained by the fact that a fast growing cell shifts resources away from recombinant protein production to creation of new biomass in preparation for cell division. In praxis a state of low growth and high production can be induced in several ways, for example by shifting pH and temperature (Trummer, Fauland et al. 2006; Trummer, Fauland et al. 2006). The influence of low temperature, i.e. 30°C, on cultivation of CHO cells was characterized on proteome level by Kaufmann and Fussenegger et al. in 1999: besides increased expression of cold-inducible proteins also changes in post-translational modification were observed, i.e. phosphorylation on tyrosine residues. In general the reduction of temperature suppresses cell growth, leading to an accumulation of cells in G1 phase. While this phenomenon of growth arrest is widely observed, the effects of low temperature on recombinant protein production

differ among various cell lines and did not uniformly result in higher specific productivity (Kaufmann et al. 1999).

### **1.2.3. Small Molecule Enhancers**

Another approach to increase specific productivity of cells is treatment with small molecule enhancers (SME) such as sodium butyrate or valproic acid (Palermo et al. 1991, Novak-Hofer et al. 2003, Wurm et al. 2008). Sodium butyrate (NaBu) has been used for more than 20 years to increase protein production. Cox and McClure showed in 1983 that HeLa cells responded to NaBu treatment with increased synthesis and secretion of alkaline phosphatase and the alpha subunit of hormone hCG (Cox et al. 1983). Palermo et al. applied NaBu to CHO cells in 1991 to increase recombinant protein production, while comparing different promoters. Recently De Leon Gatti and Wei-Shu Hu compared mouse hybridoma and CHO transcriptomes under sodium butyrate treatment and identified changes on the level of histone modification, chaperones, lipid metabolism and protein processing (De Leon Gatti et al. 2007). In 2008 Amgen scientists published results of a screening procedure to identify new classes of SMEs that affected expression of either a fluorescent reporter or a monoclonal antibody in CHO cells. They identified several new SMEs and reported an increase of up to 60% in recombinant protein titers from shake-flasks and fed-batch bioreactor experiments compared to an untreated control (Allen M J et al. 2008).

All strategies mentioned so far are confined to change the environment of mammalian cells or to external treatment of cells with SMEs. But also improvements on genomic level, for instance in DNA delivery and integration as well as host cell engineering have contributed to the high yields of modern bioprocesses:

### **1.2.4. DNA Delivery and Integration**

Design of expression vectors is crucial in order to achieve high expression of the gene of interest. Briefly, promoter and enhancer sequences of expression vectors are commonly derived from strongly expressed viral or cellular regions. Since it was found that splicing significantly increases cytoplasmic transport and translation of mRNAs, most expression vectors now include introns that are frequently located between the promoter and cDNA coding

sequence (Le Hir et al. 2003). In addition also cDNA sequence of the gene of interest may be altered in order to better match the 'codon usage' of the host cell (Makrides, 1999). Non-viral transfection methods have been largely improved in the recent past, so that transfection of expression vectors occurs with high efficiency and in a controlled manner. However, integration of the gene of interest in the host genome is still a random process although crucial for gene expression due to the so-called position effect: insertion of the transgene into heterochromatin regions of the genome results in little or no expression, while insertion into active euchromatin regions frequently yields high expression (Richards and Elgin 2002). However, also expression from euchromatin regions might be subject to gene silencing, probably due to the influence of neighboring condensed chromatin. This might be caused by hypoacetylation of histones, methylation of lysine9 on histone H3 and an increase in CpG methylation in promoter region of the recombinant DNA (Mutskov and Felsenfeld 2004). In order to avoid these negative positioning effects after random integration several strategies have been developed. For example deacetylation of histones can be blocked by addition of butyrate. Also is it possible to minimize the effect of heterochromatin by flanking the gene of interest with cis-regulatory elements such as insulators (Mutskov, Farrell et al. 2002), boundary elements, scaffold/matrix attachment regions (Girod, Zahn-Zabal et al. 2005), ubiquitous chromatin opening elements (Antoniou, Harland et al. 2003) and conserved anti-repressor elements (Kwaks, Barnett et al. 2003). Naturally the best strategy to escape positioning effects would be to target transgene integration to active parts of the genome. This could be done by including recombinases that specifically exchange DNA between a transfected plasmid and genomic DNA, if both are flanked by specific attachment regions (Wilson and Kola 2001).

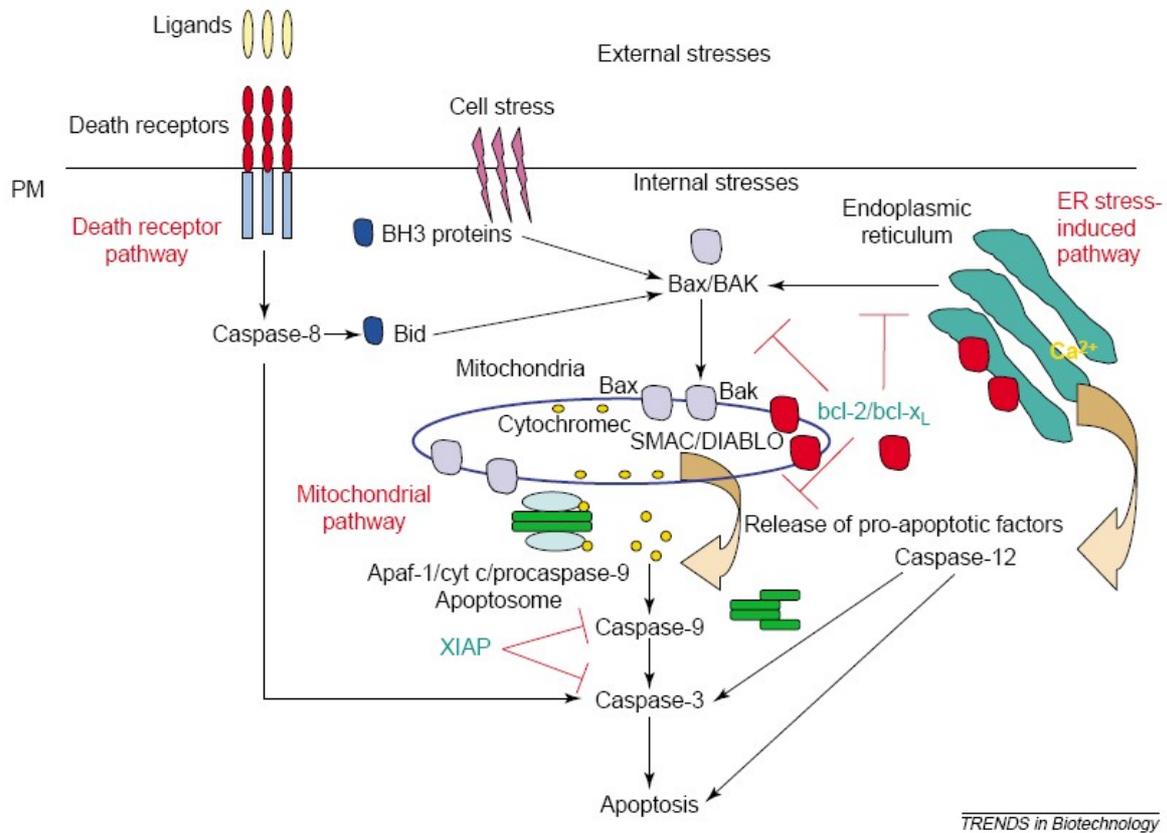
#### **1.2.5. Genetic Engineering of Host Cells**

Years of research have allowed a better understanding of several important cellular processes such as cell proliferation, programmed cell death or protein folding and secretion. Gaining insight into these molecular pathways allowed identification of key proteins, whose absence or presence can cause tremendous changes within cells. Thus, the ability to understand how key regulatory proteins control a cellular process is of great importance for the development of new strategies for cell line engineering. In general two approaches to genetic modification of

host cells are currently pursued: (i) Specific productivity of cells can be increased by improving the protein folding and secretion capacities of the cell. Since over-expression of single foldases or chaperones has not proven to successfully increase monoclonal antibody production, it was suggested that global expansion of all proteins involved in protein folding, assembly and secretion might be the way to go (Dinnis and James 2005). The challenge of simultaneous engineering of multiple genes could be addressed by applying a new molecular technology for multi-gene engineering, which Greber and Fussenegger reported in 2006: combination of transgene over-expression with knockdown of target genes was achieved by simultaneous expression of a transgene and siRNA by encoding the siRNA information in intronic sequences (Greber and Fussenegger 2007). (ii) Maintaining high cell viabilities for a longer period increases the integral of viable cells and, consequently, the volumetric productivity. This is most commonly achieved by suppressing cell death via apoptosis or by engineering metabolism so that accumulation of toxic by-products such as lactate or ammonia is avoided (Dinnis and James 2005).

#### 1.2.5.1. Apoptosis Engineering

Cell death often results in decreased product yield and eventually also impacts product quality. Consequently cell death is a major problem encountered in every bioreactor culture of mammalian cells. Two forms of cell death exist: necrosis, as the result of immediate extreme conditions that cause cells to swell and burst and apoptosis (or programmed cell death), which is a highly regulated physiological response of the cell to a specific stimulus (Alberts 2002). Morphologically, apoptosis can be characterized by cell and chromatin shrinkage, as well as shedding of membrane fragments from the cells (blebbing). Blebbing gives rise to apoptotic bodies filled with cytosolic and nuclear contents which are phagocytosed in vivo, while in vitro (e.g. during a cell culture process) they accumulate and eventually break (Arden and Betenbaugh 2004). By measuring the amount of chromatin condensation and DNA fragmentation in 180bp fragments, Daniel Wang et al. observed that about 80% of cells in a standard serum-free batch died via apoptosis (Goswami, Sinskey et al. 1999). In general there are several causes for apoptosis during a bioprocess: nutrient depletion, agitation, cell waste accumulation, hypoxia or viral infection (Arden and Betenbaugh 2004).



TRENDS in Biotechnology

**Figure 2 – Apoptosis pathway in animals:** Three pathways culminate in cell death, (i) external stress via FAS ligand and FAS receptor (death receptor) activates caspase 8; (ii) internal stress can activate pro-apoptotic molecules Bax and Bak which translocate to the outer mitochondrial membrane and cause pro-apoptotic factors such as cytochrome c or SMAC/DIABLO to enter the cytoplasm through voltage dependent anion channels (VDAC). There cytochrome c binds Apaf-1 adapter molecule and activates a caspase cascade (iii) Stress in the endoplasmic reticulum (ER) can be triggered by misfolded proteins or protein aggregation and also culminates in programmed cell death. X-linked inhibitor of apoptosis (XIAP) is a downstream inhibitor of apoptosis. Figure was taken from (Arden and Betenbaugh 2004).

Within the cell, three main pathways exist that culminate in programmed cell death (Arden and Betenbaugh 2004): (i) the mitochondrial mediated pathway; (ii) the endoplasmic reticulum (ER) stress-induced pathway; (iii) the cell-surface mediated transduction pathway. The complex network of enzymes and proteins involved in apoptosis is depicted in Figure 2. Mitochondria play a central role in apoptosis, as they contain numerous pro-apoptotic factors, such as cytochrome c, SMAC/DIABLO and apoptosis inducing factor (AIF) which are released upon arrival of internal or external stress signals. After release into the cytoplasm through voltage-dependent anion channels (VDAC) cyt c binds Apaf-1 adapter protein and together they

assemble with pro-caspase-9 (building the *apoptosome*) which results in activation of caspase 9 via auto-proteolytic cleavage. Active caspase-9, which is a protease containing a cysteine residue in its active site and which cleaves its targets at an aspartic acid residue, initiates the apoptosis cascade via caspase 3 activation. X-linked inhibitor of apoptosis (XIAP) negatively controls apoptosis on caspase level. Inhibitor of apoptosis proteins (IAP) were first identified in viruses, where they are believed to extend the life span of a cell until the virus has replicated (Alberts 2002). As can be seen in Figure 2, also the endoplasmic reticulum can trigger signals that initiate apoptosis. This makes sense because ER-malfunction, for example due to errors in protein glycosylation, expression, folding and transport to the Golgi apparatus – commonly referred to as the *unfolded protein response* - can have severe adverse effects on neighboring cells due to secretion defective proteins.

One successful strategy to delay the event of apoptosis in a serum-free cell culture is to add specific supplements such as nutrients, or anti-apoptotic proteins or peptides to the media (Sunstrom, Gay et al. 2000; Zanghi, Renner et al. 2000). This approach is rounded up by genetically engineering the apoptosis pathway within cells, where members of the Bcl-2 protein family are prominent targets. Bcl-2 family members share so-called Bcl-2 homology (BH) domains, and, depending on the set of BH domains they exhibit either pro-apoptotic or anti-apoptotic functions. Bcl-2 and Bcl-x<sub>L</sub> are located in the outer membrane of mitochondria and both contain BH domains 1 to 4. They have been described to exert anti-apoptotic effects by inhibiting the release of pro-apoptotic molecules from the mitochondria, either by binding pro-apoptotic bcl-2 family members, Bax and Bak, or by maintaining the mitochondria membrane potential (Harris and Thompson 2000; Kroemer 2003; Scorrano and Korsmeyer 2003). Overexpression of Bcl-x<sub>L</sub> and especially Bcl-2 is the most frequently chosen strategy for apoptosis engineering. Several studies have shown that their overexpression results in increased viabilities in CHO, NS0 or BHK cells under adverse conditions such as nutrient depletion or exposure to toxins or viral infection (Chao, Linette et al. 1995; Tey, Singh et al. 2000; Tey, Singh et al. 2000). The idea of overexpressing Bcl-2 and Bcl-x<sub>L</sub> has been taken to the next level by engineering these proteins in order to become more resistant to caspase cleavage. This makes sense because in the course of apoptosis Bcl-2 is processed by caspases, leading to

its conversion from anti-apoptotic to pro-apoptotic. By deleting the cleavage site, which is a loop structure between BH domains 3 and 4, Bcl-2 $\Delta$  becomes more resistant to degradation. CHO cells overexpressing Bcl-2 $\Delta$  perform better compared to wild-type Bcl-2 overexpression upon exposure to multiple insults (Figueroa, Sauerwald et al. 2001; Figueroa, Sauerwald et al. 2003). Also, Kim and Lee showed in 2000 that it makes sense to combine the use of anti-apoptotic engineered cell lines with SME treatment, which on its own has been shown to have adverse effects on cell viability. This way increased viabilities as well as a high STY were achieved (Kim and Lee 2000). Besides engineering of bcl-2 family member proteins also genetic engineering of XIAP (X-linked inhibitor of apoptosis), caspase-9 and caspase-3 as well as the introduction of viral proteins, which maintain the mitochondrial membrane potential and target bcl-2 family proteins, have been tested for apoptosis engineering (Arden and Betenbaugh 2004; Mohan, Kim et al. 2008). What is more, overexpression of two anti-apoptotic genes, Aven and E1B-19K, in Baby Hamster Kidney cells (BHK) gave rise to a transformed cell line with increased performance in recombinant protein production (Figueroa, Ailor et al. 2007; Nivitchanyong, Martinez et al. 2007).

However, overexpression of transgenes in mammalian cells, especially when it comes to multi-gene engineering has the disadvantage of increasing the metabolic burden on a cell. This is due to the massive overexpression of proteins on top of the enhanced expression of the actual recombinant protein of interest (Yallop and Svendsen 2001; Yallop, Norby et al. 2003). Thus, Müller and Grillari proposed a new approach to cell line engineering that offers global regulation of cellular pathways while adding less extra burden to a cell's metabolism: microRNA (Muller, Katinger et al. 2008).

### **1.3. Engineering Host Cells on microRNA Level**

#### **1.3.1. History of microRNAs**

Small non-coding RNA molecules referred to as microRNAs (miRNAs) were first discovered in the nematode *Caenorhabditis elegans*, where the *lin-4* gene was shown to encode not a protein, but a small RNA species that exerted regulatory function on *lin-14*, a protein involved in development of *C. elegans* (Lee, Feinbaum et al. 1993). It was found that a conserved region

in the 3'-untranslated-region (3'UTR) of *lin-14* contained partially complementary sequences to *lin-4* small RNA – the key to the regulatory function of miRNAs (Wightman, Ha et al. 1993). During further studies miRNAs were identified in every metazoan organism analyzed so far, where they seem to not only regulate timing of development, like for example *lin-4*, but also other biological processes such as cell death, cell proliferation or cell metabolism (Ambros 2004).

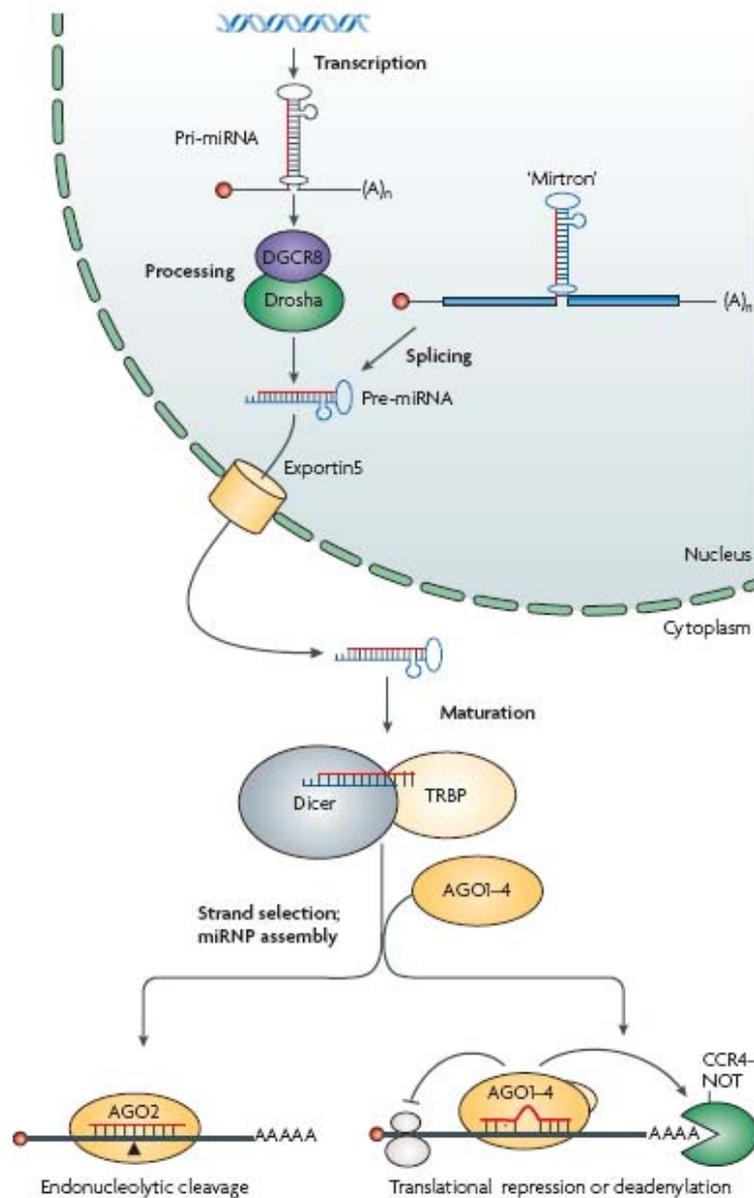
### **1.3.2. MicroRNA Genomics and Biogenesis**

Today it is known that miRNAs are 20-nt to 22-nt long regulatory non-coding RNAs that are expressed in multicellular organisms as well as unicellular organisms and large DNA viruses. In archaee as well as eubacteria no miRNAs have been identified so far (Bartel 2004; Filipowicz, Bhattacharyya et al. 2008). In general miRNA sequences are diverse, with the exception of a Uracil-residue which frequently appears at the 5' end of miRNAs (Lewis, Burge et al. 2005). Based on sequence homology or organization in genomic clusters miRNAs can be grouped into families, of which many seem to be highly conserved across species (Bartel 2004). Identification and prediction of miRNA genes is achieved by applying what Lee and Wightman had already found out for *lin-4*: mature 22-nt long miRNA sequences are part of sequences that form imperfect stem-loop structures, with the mature miRNA located within one arm of this stem-loop (Lee, Feinbaum et al. 1993; Wightman, Ha et al. 1993). When again looking at cross-species conservation, strong sequence conservation can be found in the stems of these miRNA hairpins, while lower conservation is found in the loop structure and especially in stem flanking regions. Using this information and bioinformatic tools such as phylogenetic shadowing, it was possible to predict miRNA genes by comparing human/mouse and human/rat conservation and looking for certain conservation patterns. The outcome was that the human genome might contain as many as 1000 miRNA genes, which would make miRNAs account for 2 – 3% of all genes in the human genome (Berezikov, Guryev et al. 2005).

Genomic analysis has further shown that miRNA genes (including regulatory elements) are either located far away from previously annotated regions, or that they originate from intronic or exonic regions within transcriptional units (TU). Only recently it turned out that the majority

of miRNA genes (70%) are contained within introns or exons of TUs (Kim and Nam 2006). Within this group three possible arrangements with different frequencies exist: (i) intronic miRNA in protein-coding TU (61%); (ii) Intronic miRNA in non-coding TU (18%); (iii) exonic miRNA in non-coding TU (20%) (Kim and Nam 2006; Ruby, Jan et al. 2007). Furthermore about 50% of all microRNA genes are organized in clusters, from which they are transcribed as polycistronic primary transcripts. It has been suggested that clusters of related miRNAs are the consequence of gene duplication events, while clusters of unrelated miRNAs (in terms of sequence) might be related as far as that they target mRNAs of the same cellular pathway (Kim and Nam 2006).

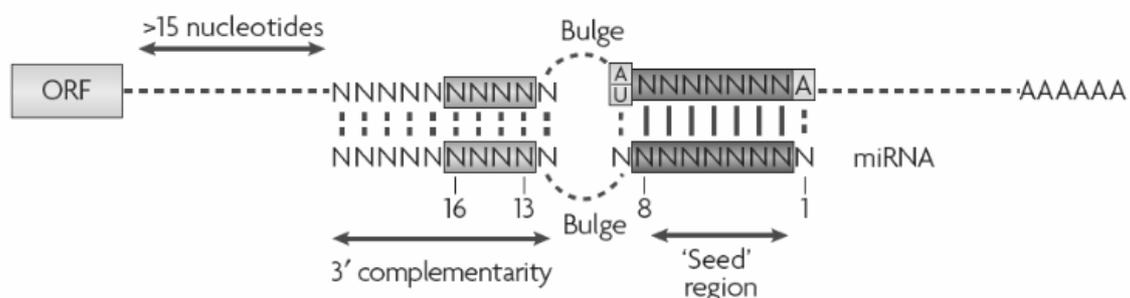
Primary miRNA transcripts (pri-miRNAs) are often greater than 1kb in length and, as already mentioned, contain a local fold-back structure referred to as stem loop or hairpin. Several findings indicate that pri-miRNAs are RNA Polymerase II transcripts: they carry a five-prime 7-methyl guanosine cap as well as a three-prime poly(A) tail. Also, miRNA promoter studies have identified typical Pol II elements such as a TATA Box upstream of miRNA transcription sites (Houbaviy, Dennis et al. 2005). Upon transcription stem-loop structures within the pri-miRNAs are co-transcriptionally processed in the nucleus (Morlando, Ballarino et al. 2008) by an RNaseIII type endonuclease called Drosha, yielding an approximately 70-nt to 80-nt long precursor-miRNA or pre-miRNA. Drosha is a large (~160 kDa) and highly conserved protein that requires a cofactor, DGCR8 in humans (Pasha in *Drosophila* and *C. elegans*). Due to two double-stranded RNA-binding domains (dsRBD) in DGCR8, this cofactor is believed to assist Drosha in pri-miRNA recognition. In the next step Exportin5 (Exp5), a Ran-guanosine triphosphate dependent transporter, exports the precursor miRNA to the cytoplasm. There another RNaseIII type endonuclease, Dicer, binds the precursor miRNA and cleaves it into a miRNA duplex with 2-nt overhangs on both 3' ends. Similar to Drosha, also Dicer is a large (~200 kDa) and highly conserved protein, which associates with a dsRBD-containing cofactor, R3D1 in *Drosophila* and TRBP in humans (Bartel 2004; Kim and Nam 2006; Filipowicz, Bhattacharyya et al. 2008). Only one strand of the ~22nt-duplex intermediate becomes the mature miRNA, presumably the strand with less stable hydrogen bonding at its 5' end. This finding might also answer the high frequency of U-residues on mature miRNA 5' ends.



**Figure 3 – Biogenesis of microRNAs:** Transcription in the nucleus either from intergenic or intronic regions yield a precursor-miRNA (pre-miRNA) that is exported to the cytoplasm through Exportin5. Dicer TypeIII endonuclease processes pre-miRNA to a ~22nt duplex with 3'overhangs. The mature miRNA strand is incorporated in a ribonucleoprotein complex (miRNP) which contains proteins from the Argonaute (AGO) protein family. Depending on the quality of base-pairing to its target, microRNAs either exhibit endonucleolytic cleavage or translational repression. Figure taken from (Filipowicz, Bhattacharyya et al. 2008).

### 1.3.3. Function of miRNA as post-transcriptional Regulators

After mature microRNAs are processed from their pre-miRNA precursors they are assembled into a ribonucleoprotein complex, referred to as miRISC (miRNA induced silencing complex) or miRNP (miRNA ribonucleoprotein), whose major components are proteins of the Argonaute (AGO) family. In mammals four AGO proteins are involved, AGO1 to AGO4. Other proteins contained in miRISCs are regulatory factors or effectors that mediate the inhibitory function of miRISCs (Peters and Meister 2007). Once incorporated into this protein complex, the miRNA directs the silencing complex to a messenger RNA that harbors a complementary sequence. While in plants miRNAs base-pairing to mRNA is perfect in most cases, leading to mRNA degradation (Jones-Rhoades, Bartel et al. 2006), miRNAs in mammals tend to imperfectly match their target sites in the 3'UTR of mRNAs resulting in repression of translation. A set of requirements for vertebrate miRNA target recognition was proposed based on experimental and bioinformatic analysis and is shown in Figure 4 below (Brennecke, Stark et al. 2005; Lewis, Burge et al. 2005; Grimson, Farh et al. 2007; Filipowicz, Bhattacharyya et al. 2008).



**Figure 4 – Requirements for miRNA target recognition upon imperfect base-pairing:** 1) Perfect match from base-pair 2-8 (seed region). 2) Bulge mismatch following the seed region. 3) Complementary 3' from Bulge, which allows some mismatches. Figure taken from (Filipowicz, Bhattacharyya et al. 2008).

Briefly, the most important requirement is perfect base pairing of the miRNA nucleotides 2-8, also referred to as the seed region which is located at the 5' end. Also required for functional miRNA target recognition are bulges or mismatches in the central region of the miRNA:mRNA duplex, followed by 'reasonable' complementarity in the 3' half of the miRNA.

#### **1.3.4. Relevance of microRNAs as Key Regulators in Animal Cells**

MicroRNAs have been shown to act as regulators within eukaryotic cells via post-transcriptional regulation of target mRNAs (Bartel 2004). Their global impact on cells was demonstrated by transfecting single miRNAs into HeLa cells and profiling mRNA levels (Lim, Lau et al. 2005). Lim et al. found that transfection of miR-1 or miR-124 caused a major shift in the mRNA profile, by downregulating about 100 mRNAs and shifting the expression profile towards muscle cells or towards brain cells, respectively. After miRNA regulation being first related to developmental processes in nematodes and flies this study emphasized the importance of miRNAs for tissue specific gene regulation.

Other studies underlined the importance of miRNAs for development and progression of diseases such as cancer (Hammond 2006) and diabetes (Hennessy and O'Driscoll 2008). The importance of miRNAs for cancer is based on their regulation of cell proliferation and apoptosis (Cheng, Byrom et al. 2005; Cimmino, Calin et al. 2005); while miRNA were associated with diabetes due to their involvement in metabolism and secretion (Kruzfeldt and Stoffel 2006). Obviously, ways of regulating these cellular programs are of interest to the bioprocess community as well and investigation of miRNA regulation in animal cells, i.e. CHO, is the logical consequence.

In 2007 Gammell et al. published the first miRNA sequence identified in Chinese Hamster Ovary cells (cgr-miR-21) while investigating CHO microRNAs using microarray, quantitative PCR and cloning technology (Gammell, Barron et al. 2007). MiRNA microarrays used by Gammell et al. were spotted using miRNA sequences for human, mouse and rat since no public sequence information had been available for CHO at that time. Gammell et al. identified 26 conserved miRNAs with differential expression between high and low temperature (37°C vs. 31°C), mimicking a biphasic cultivation. Using quantitative PCR technology they were able to confirm upregulation of two microRNAs under cold-shock conditions, miR-21 and miR-24. Both microRNAs also showed up as differentially expressed when early (72h) and late (144h) culture stages at 37°C were compared. Finally miR-21 was cloned from CHO using primers based on human, mouse and rat genomic information of pre-miR-21 flanking regions. Finally, sequence

information of cgr-miR-21 was published at Sanger miRBase as the first Chinese Hamster Ovary miRNA.

### **1.3.5. MicroRNA-21**

Since 2005 microRNA-21 has been the subject of 18 published studies on cancer and cancer progression. Interestingly it was found to be the only commonly overexpressed miRNA in solid tumor tissues, i.e. breast, brain, colon, lung, liver, prostate and stomach, leading to the term 'oncomir' (Hammond 2006). The impact that increased levels of microRNA-21 in cancer tissues have seem to be manifold: in glioblastoma cells miR-21 was reported to be highly expressed; upon specific miR-21 knockdown increased occurrence of apoptosis was observed due to higher caspase activity. Hence miR-21 was attributed an anti-apoptotic activity. However, the exact interactions of miR-21 that inhibit apoptosis are not completely clear, but Sun and Peng suggested it to be the result of a large number of miR-21 interactions, for example with PTEN, SOX5 or even Bcl-2 (Chan, Krichevsky et al. 2005; Chen, Liu et al. 2008).

Interaction of miR-21 with tumor suppressor gene PDCD4 (programmed cell death 4) has been identified in colorectal cancer cell lines, where this interaction was suggested to stimulate invasion, intravasation and metastasis (Asangani, Rasheed et al. 2008). In MCF-7 cells, which is a commonly used early breast cancer cell line with high levels of PDCD4 and relative low levels of miR-21, neoplastic transformation has been shown to be influenced by mature miR-21 levels (Frankel, Christoffersen et al. 2008). PDCD4 is currently the most frequently cited target of miR-21 which is why it makes sense to take a closer look on this protein: about a decade ago the mouse mRNA 'MA-3' was identified to be upregulated during apoptosis in several cell lines such as thymocytes, T cells and B cells (Shibahara, Asano et al. 1995). In the following years homologous genes in human, chicken and rat were identified and for the sake of clarity were termed PDCD4 referring to *programmed cell death 4*, or *neoplastic transformation inhibitor protein* (Lankat-Buttgereit and Goke 2003). PDCD4 is a 485 amino acid long protein that contains two MA-3 domains which are thought to be responsible for its function. In 2008 it was shown that these MA-3 domains bind MA-3 domains on eukaryotic initiation factor 4-A (eIF4A), thus inhibiting its helicase activity and interfering with cap-dependent translation (Suzuki, Garces et al. 2008). Using JB6 cells, the function of PDCD4 as a tumor suppressor was described.

Briefly, JB6 cells are either promotion-sensitive (P+) or promotion-resistant (P-), meaning that the cells are either likely to undergo neoplastic transformation or not. In contrast to P+ cells, P- cells were found to highly express PDCD4, which suggested a role for PDCD4 as tumor suppressor. Also inhibition of PDCD4 in P- cells via stable transfection of antisense PDCD4 resulted in a phenotype shift to P+, thus knockdown of PDCD4 promoted neoplastic transformation. When PDCD4 was overexpressed in P+ cells, and neoplastic transformation was induced by adding TPA (Tetradecanoylphorbol acetate), no transformation was observed which again demonstrated its function as a tumor suppressor (Cmarik, Min et al. 1999; Yang, Jansen et al. 2001; Yang, Knies et al. 2003). Recently it has been tried to explain these effects on molecular level by down-regulation of MAP4K1 through PDCD4, which inhibits certain events that would drive invasion such as MAPK85 activation and consequent JUN-dependent transcription (Yang, Matthews et al. 2006).

Besides regulation of transformation, invasion and metastasis miR-21 has also been related to tumor growth in breast cancer tissue. Si et al. showed that inhibition of intracellular miR-21 levels via anti-miR-21 transfection, resulted in reduced cell growth in vitro as well as reduced tumor-growth in the in a xenograft mouse model (Si, Zhu et al. 2007). Thus, miR-21 was described as an interesting drug target for cancer therapies. On the other hand it was shown that some cell models seemed to behave quite differently upon miR-21 inhibition or overexpression. For example Cheng et al. screened the effect of several miRNA inhibitors on cell growth and apoptosis of HeLa cells. Here anti-miR-21 transfection resulted in increased growth compared with a non-transfected control, while no effect of miR-21 knockdown on apoptosis (by measuring caspase-3 activity) was observed (Cheng, Byrom et al. 2005). For apoptosis this might be a cause of choosing an improper readout, but cell proliferation was clearly affected differently in HeLa cells than in cancer cells.

Besides miR-21 interaction with tumor suppressor PDCD4, interaction has also been confirmed with several other tumor suppressors such as PTEN, maspin and Tropomyosin1 (Zhu, Si et al. 2007; Zhu, Wu et al. 2008).

With these interactions being at the moment confined to specific cancer cell lines it is not clear how Chinese Hamster Ovary cells, as a transformed cell line, will respond to miR-21



In the first step, applying small RNA deep sequencing technology will allow identification of conserved miRNAs in CHO by comparing sequences with publicly available miRNA sequence information, primarily human, mouse and rat. In addition to identifying conserved miRNA sequences, Illumina/Solexa Sequencing data can also be used for novel (CHO species specific) miRNA prediction by applying RNA folding algorithms to unannotated small RNA reads. Knowledge about CHO miRNA sequences combined with information about target interactions will allow to specifically look for miRNAs that act as global regulators within cellular pathways relevant for bioprocessing, such as apoptosis or cell proliferation. In the next step these assumptions need to be verified by subjecting selected miRNAs to expression and inhibition studies in a CHO production cell line.

Being one of the best studied microRNAs so far, microRNA-21 has been suggested to play important roles in cancer development and progression via regulation of transformation, invasion, metastasis, cell growth and apoptosis. Thus for a start it seems promising to study the effect of inhibition as well as overexpression of miRNA-21 in CHO cells under various conditions. These conditions either need to be relevant for bioprocesses such as batch cultivation, or need to resemble problems that might occur in bioreactors such as nutrient limitation or oxidative stress.

Eventually knowledge created in this work can be applied to develop a novel - microRNA-engineered - CHO cell line, whose performance in a bioreactor needs to be assessed and compared with other engineered recombinant cell lines. Also the comparison of miRNA expression patterns of various CHO cell lines or of the same cell line under various conditions (such as temperature shift or sodium butyrate treatment) is going to help elucidate the importance of miRNA regulation in Chinese Hamster cells.

In the end, the results of these studies will help decide whether miRNA technology can become a valuable tool for cell line development and screening for the bioprocessing community.

### **3. Material and Methods**

#### **3.1. Cell Culture**

##### **3.1.1. Cell Lines and Media Composition**

Three recombinant CHO cell lines were used in this study, CHO-EpoFc, CHO-Aven/E1B19k (CHO-Aven) and CHO-MDJ-84.

CHO-EpoFc cells were kindly provided by Polymun Scientific (Vienna, Austria) and, in brief, were derived from the dhfr-deficient parental cell line CHO-DUKX-B11 (ATCC CRL-9096) and subsequently adapted to grow in suspension and in the absence of serum. Cells are stably transfected with a homodimeric Epo-Fc fusion protein that was generated by inserting human IgG<sub>1</sub>-Fc gene downstream to the Epo coding region (Trummer, Fauland et al. 2006). After revitalization from a research cell bank, CHO-EpoFc cells were cultivated in Dulbecco's Modified Eagle's Medium DMEM/Ham's F12 (1:1), supplemented with 4mM L-Glutamine, 0.25% ultra-filtrated soy peptone, 0.1% Pluronic F68 and in-house developed 1x protein-free additive (proprietary formulation). Selective pressure was maintained by addition of 0.096  $\mu$ M MTX. CHO-EpoFc cells were routinely cultivated in 125 ml Spinner Flasks (Techne, Staffordshire, UK) at a working volume of 60 ml and a stirrer speed of 50 rpm; 50 ml of filtered (0.22 $\mu$ m) carbon dioxide were added to the headspace in order to buffer the pH.

Cells were subcultured every 3 or 4 days and seeded at a cell density of  $1.5 \times 10^5$  cells/ml into fresh media.

CHO-Aven, which was developed in Michael Betenbaugh's lab, is derived from the dhfr-deficient parental cell line DG44 and was engineered by overexpressing two viral anti-apoptotic genes, Aven and E1B19k (Figuroa, Ailor et al. 2007). CHO-Aven cells are adapted to grow in suspension and in the absence of serum in commercial CHO-S-SFM II media (Invitrogen, Carlsbad, CA). Cells were cultured in T25, T75 and T150 culture flasks at 37°C in a 5% CO<sub>2</sub> atmosphere, and were subcultured every 3 or 4 days to a cell density of  $3 \times 10^5$  cells/ml.

CHO-MDJ-84 cells, which were retrieved from Florian Wurm's lab, were grown in the presence of 5% fetal bovine serum using Dulbecco's Modified Eagle's Medium DMEM/Ham's F12 (1:1) at 37°C and 5% CO<sub>2</sub> atmosphere. Cells were passaged every 4 days using a split ratio of 1:4. CHO-

MDJ-84 cells produce a monoclonal antibody but have not yet gone through the process of gene amplification.

### 3.1.2. Cell Culture Parameters

In order to describe cell growth under batch conditions specific growth rate  $\mu$  as  $\text{day}^{-1}$  was calculated. Therefore the following equation was derived from the basic equation for the change of biomass in a reactor under batch conditions ( $dx/dt = \mu \cdot x$ ), where  $x_2$  and  $x_1$  correspond to cell concentrations determined at timepoints  $t_2$  and  $t_1$ :

$$\mu = \frac{\ln(x_2) - \ln(x_1)}{t_2 - t_1}$$

For description of cell productivity the specific production rate  $q_p$  was calculated. Therefore the integral of viable cells (IVC), also referred to as cumulative viable cell days (CD) had to be calculated, since the accumulation of secreted protein is not directly related to the course of viable cell concentration, but to the area (integral) below this curve. The following equation was used for calculation of cumulative cell days using viable cell concentration  $x_{i+1}$  and  $x_i$  within a time interval  $t_1$  to  $t_n$  with discrete timepoints  $t_i$ :

$$IVC = \sum_{i=1}^n \frac{x_{i+1} - x_i}{\mu}$$

The unit of CD is cells\*days. Specific productivity  $q_p$  was then calculated as product accumulation over a certain period divided by the CD of the same period. Variables  $c_n$  and  $c_0$  correspond to EpoFc concentration at the end and beginning of a cultivation and CD to the integral of viable cells:

$$qP = \frac{c_n - c_0}{CD}$$

The unit of specific productivity  $q_p$  used in this study was attomol per cell per day (amol/cell/day). Metabolic uptake rates were calculated analogical.

The volumetric productivity or space-time-yield (STY) describes the accumulation of secreted product in relation to the process time, and was calculated using the following equation, where  $c_n$  and  $c_0$  again describe the EpoFc protein-concentration at the end and beginning of the cultivation, respectively, and  $t_n$  and  $t_0$  the corresponding timepoints:

$$STY = \frac{c_n - c_0}{t_n - t_0}$$

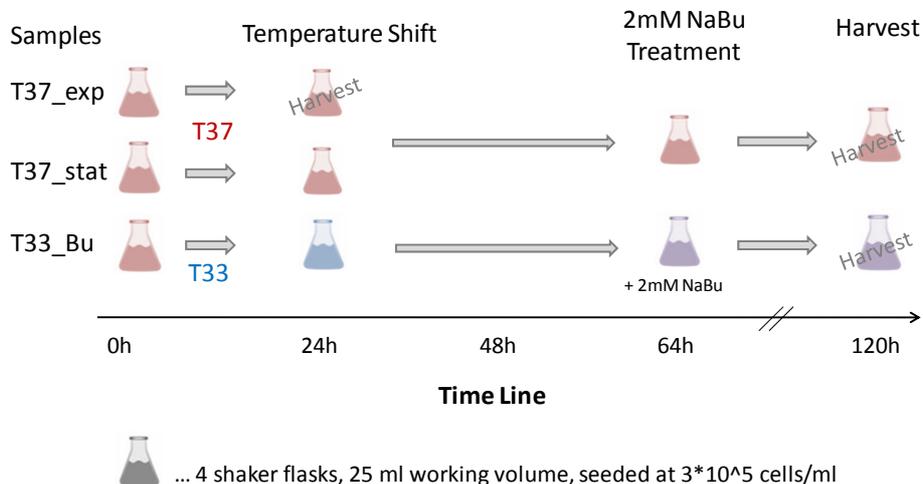
All equations listed above were used according to directions and explanations from a publication by Trummer et al. (Trummer, Fauland et al. 2006).

### **3.1.3. Temperature Shift and Sodium Butyrate Treatment**

The use of sodium butyrate (NaBu) is widespread in mammalian cell culture and has been shown to positively influence recombinant protein production in CHO cells (Palermo, DeGraaf et al. 1991). Therefore common cultivation strategies that combine an initial phase of fast cell growth with a second phase of growth inhibition and high productivity, use NaBu to further enhance protein production during stationary phase .

CHO-Aven cells were expanded to 10 shaker flasks at a density of  $3.0 \cdot 10^5$  cells/ml, in CHO-S-SFM II media. Working volume was 25 ml and rotation speed was set to 100 rpm. Figure 6 below shows a scheme the time course of the cultivation and indicates treatments and total RNA harvests. Cells were incubated at 37°C in humidified air containing 5% CO<sub>2</sub> for 24h or until cell density reached  $5 \cdot 10^5$  cells/ml. At this time 4 shaker flasks were harvested using Trizol (Invitrogen, Carlsbad CA). Total RNA extracted from these cells represented exponential growth phase samples. Also after 24h, three shaker flasks were transferred to an incubator at 33°C and a 5% CO<sub>2</sub> atmosphere while the remaining three shaker flasks were maintained at 37°C. Following 42 hours of incubation at 33°C, sodium butyrate was added to all 3 of the flasks to a working concentration of 2 mM. Cells were cultivated for an additional 60h before harvested in

Trizol. Overall the cultivation lasted for 120h and samples were taken every day and analyzed for cell density and cell viability.



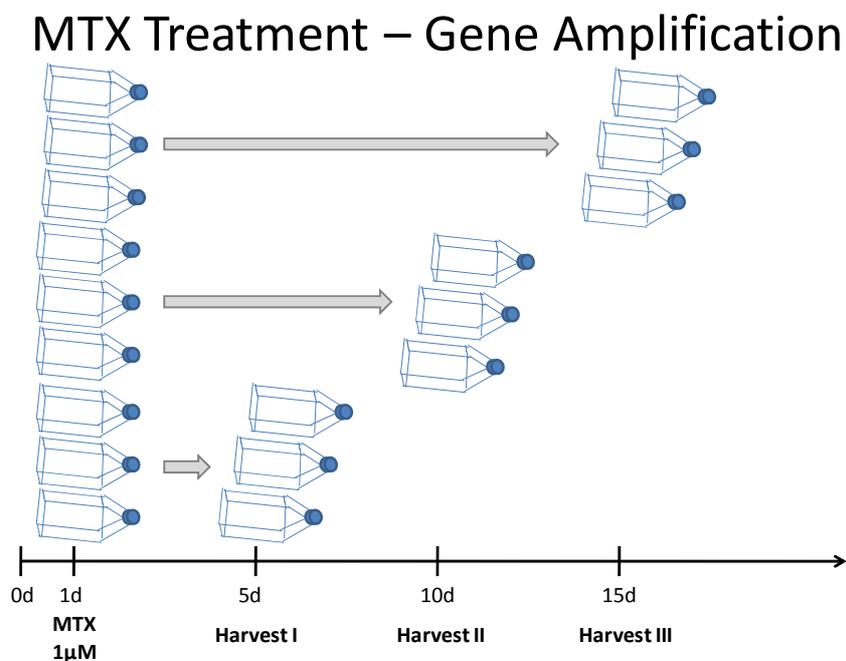
**Figure 6 – Experiment layout for low temperature and sodium butyrate treatment of CHO-Aven cells.**

### 3.1.4. Methotrexate Gene Amplification

In order to make the process of gene amplification more effective it has become increasingly common among Biotech companies to perform single step methotrexate (MTX) gene amplifications using elevated MTX concentrations. In brief, this amplification strategy relies upon the effect of MTX as a substrate analogue for Dihydrofolatereductase (dhfr). Sophisticated addition of MTX inhibits growth of cell with low dhfr expression levels but enriches the culture for cells that have gone through gene amplification of dhfr and therefore have a growth advantage. Thus, a gene of interest that had been introduced based on the dhfr selection system was also amplified and consequently cells exhibit higher expression of the recombinant protein.

For this study,  $1\mu\text{M}$  MTX working concentration for single step amplification was chosen. CHO-MDJ-84 cells were expanded to fourteen T150 flasks, subcultured using a split ratio of 1:4 and incubated for 24h at  $37^\circ\text{C}$  in a 5%  $\text{CO}_2$  atmosphere. At this time-point, when cells had attached to the culture flask surface and had entered exponential growth phase,  $1\mu\text{M}$  MTX was added to 10 out of 14 flasks. As a negative control 4 untreated T150 flasks were harvested after another 24h, while still growing exponentially. During MTX treatment cells were harvested every 5 days

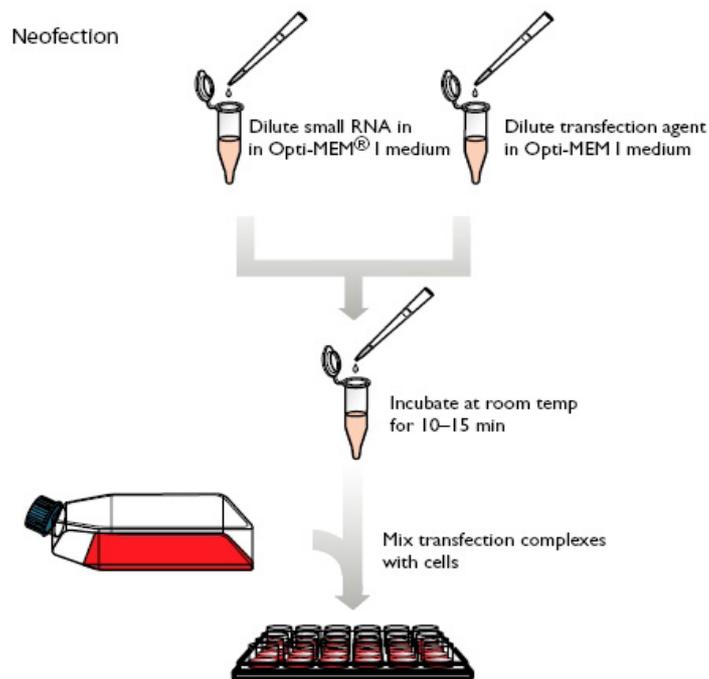
for 15 days in Trizol. Fresh media and MTX were added every 4 days. Figure 7 overviews the MTX gene amplification assay and indicates time-points of total RNA harvests from treated cells.



**Figure 7 – Experiment layout for single-step methotrexate gene amplification.**

### 3.1.5. MicroRNA Precursor Transfection

For overexpression and inhibition of microRNAs, Ambion reverse transfection technology (Neofection, Ambion Inc., Austin, TX) was used. Neofection is a novel and less time-consuming lipid-based non-viral transfection method, since cells do not have to be seeded 24h prior to transfection. In general, lipid-based transfection methods are based on combining nucleic acids with a lipid (the so-called transfection agent) so that a nucleic-acid/lipid complex is formed. This complex then attaches to the cell membrane of animal cells and enters the cell via endocytosis. Special lipid compositions ensure that the endocytosed vesicle is degraded before its cargo is transported to the lysosomes for degradation.



**Figure 8 – Overview of pre-miRNA reverse transfection protocol.**

For this study, microRNA-21 was retrieved from Ambion Inc. in the form of precursor-miRNAs. For inhibition of endogenous miRNAs, antagomir molecules (anti-miR-21) were derived that specifically bind and down-regulate complementary mature miR-21 (Cheng, Byrom et al. 2005). For each transfection two negative controls were included: (i) Ambion negative-control pre-miRNAs (nc-pre-miR) which have the same structure as functional precursors but no targets and (ii) mock transfections where instead of RNA medium was mixed with the transfection agent. Two transfection agents were tested in this study, siPORT Amine and siPORT Neo<sup>FX</sup> (both from Ambion Inc., Austin, TX). Before use, dried precursors were centrifuged and diluted in nuclease-free water to a final concentration of 6.25  $\mu\text{M}$ . MicroRNA transfections were performed in CHO-EpoFc cells, in 6-well plate format, with an initial cell density of  $1 \times 10^5$  cells/ml according to the flow chart in Figure 8. Therefore, in the first step cells were counted and diluted to the correct cell density. Prior to transfection the lipid agent was pre-warmed to room temperature and 5  $\mu\text{L}$  of were then mixed with 95  $\mu\text{L}$  of Opti-MEM II media (Gibco, United Kingdom). In the next step 12  $\mu\text{L}$  of miRNA precursor or antagomir stock were diluted with 88  $\mu\text{L}$  nuclease free water, added to the transfection agent and gently mixed. The mix was incubated for 10 minutes

at room temperature to allow formation of RNA/lipid transfection complexes, and then transferred to a well in a 6-well plate. Finally 2300µl of cell suspension were added to the well, gently mixed and incubated at 37°C and 7% CO<sub>2</sub>. All steps were carried out using sterile and nuclease free microcentrifuge tubes (Sarstedt, Germany).

## 3.2. Analytical Methods

### 3.2.1. Cell Number and Cell Viability

#### 3.2.1.1. Hemocytometer

Hemocytometers were originally developed for counting red blood cells, but are nowadays frequently used in mammalian cell culture for analyzing cell numbers and viabilities. A hemocytometer consists of a thick glass microscope slide with a rectangular indentation that creates two chambers. Each chamber has a laser engraved grid of exact known size and, after closing with a cover slide, of exact volume. Number of cells counted in the grid combined with information about the grid size and chamber depth of 0.1mm is used to calculate the cell density of a sample as cells per milliliter (cells/ml).

For routine cell culture analysis cells were counted in a total of 8 squares (1mm<sup>2</sup> size). Thus, with a depth of 0.1mm one square corresponds to a volume of 0.1mm<sup>3</sup> or 0.1µL. In order to calculate cells per milliliter, the initial count has to be divided by the number of squares counted (i.e. 8), and multiplied by 10000. When cells were previously stained with Trypan Blue, a diazo dye which only protrudes membranes of dead cells, also information about cell viability can be gained. For calculation of cell count the dilution factor (i.e. 1.2) has to be taken into account:

$$\text{Total Cell Count} \left( \frac{\text{cells}}{\text{ml}} \right) = \frac{\text{Number of total cells counted}}{\text{Squares counted}} * 10000 * 1.2$$

$$\text{Viability (\%)} = \frac{\text{Number of colorless cells}}{\text{Number of total cells}} * 100$$

$$\text{Viable Cell Count} \left( \frac{\text{cells}}{\text{ml}} \right) = \text{Total Cell Count} * \frac{\text{Viability}}{100}$$

### 3.2.1.2. Automated Counting of Cell Number – Beckman Multisizer 3™

In order to minimize errors and increase throughput, automated cell counting devices have been developed, the Beckman MultiSizer 3™ Coulter Counter being one of them (Beckman Coulter Inc., Fullerton, CA). This device consists of a glass aperture that separates two electrodes. Between these two electrodes an electric current flows. Cells or nuclei suspended in an electrolyte enter the aperture and, while passing the “sensing zone” between the electrodes, create a pulse that is directly proportional to cell size or nucleus size, respectively.

For routine analysis using Multisizer 3™, 2 ml of cell suspension were centrifuged at > 170g for 10 minutes. Supernatant was carefully discarded and the cell pellets were re-suspended in 1 ml of a TRIS/Citric Acid cell lysis buffer (T/C buffer). Cells were mixed thoroughly and incubated at room temperature for at least 2 hours. By this time cells had been completely lysed and only cell nuclei remained intact. Depending on cell density (e.g. known from prior hemocytometer analysis), a specific volume, usually in the range of 50µl to 200µl, of cells was diluted in 9 ml of Coulter Buffer (Beckman Coulter Inc., Fullerton, CA) and drawn into the aperture. Total cell number (since no distinction between viable and live cells can be made) was counted using the following equation:

$$\text{Total Cell Count} = (\bar{x} * 2 * \frac{9 + a}{a}) / b$$

$\bar{x}$  ... Average of two cell count results

$a$  ... Volume of lysed cell suspension applied to Beckman Coulter

$b$  ... Volume of cell suspension originally taken

In order to calculate viable cell concentration an additional method for viability analysis had to be performed on the same sample.

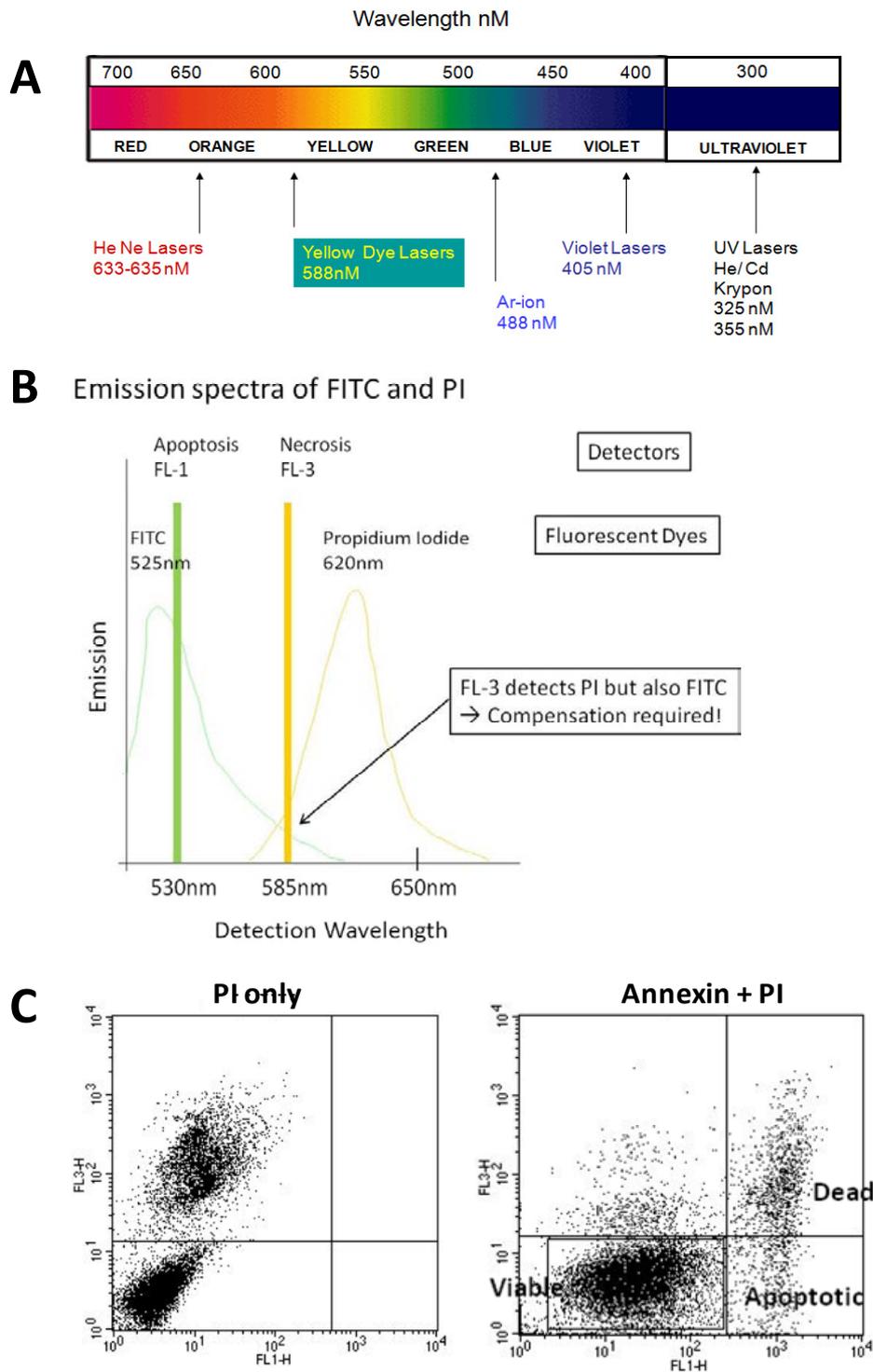
### 3.2.2. Flow Cytometry

Flow cytometry technology is commonly used for counting, examining and sorting microscopic particles, often cells. Cells suspended in a hydro-dynamically focused stream of fluid pass optical and/or electrical detectors which detect light that has been emitted from a laser source; often several lasers with different wavelengths are installed. When a single cell passes through the laser light, the light becomes scattered in several ways. Detectors in line with the light beam measure forward scattered (FSC) light, detectors perpendicular to the light measure side scattered (SSC) light. FSC yields information about cell size, whereas SSC corresponds to intracellular characteristics, such as size of nucleus, cytoplasm and granulation of cells passing through the light. Also, fluorescent light detectors are installed, so that light coming from fluorescently labeled cells can be detected. By specifically labeling certain cell populations with fluorescent dyes, flow cytometry can be used to sort (separate) populations from the original sample, which is referred to as fluorescent-activated cell sorting (FACS). Applying several fluorescent labels to a cell is possible if they emit light at different wavelengths.

For this study, flow cytometry was predominately used for analyzing cell viability, apoptosis and cell death as well as for measuring transfection efficiency of fluorescently labeled microRNAs. Therefore FACSCalibur (Becton Dickinson, Franklin Lakes, N.J.) device was used. Fluorescein (FITC)-labeled molecules were excited using a 15-mW, 488-nm air-cooled argon ion laser, and FITC emission was measured using bandpass filter for 530-nm (FL-1). Propidium iodide was excited at 635-nm, using a red diode laser, and PI emission was detected with a bandpass filter at 585-nm (FL-3). A total of 10000 cells were measured for every sample.

#### 3.2.2.1. Dual-Dye Measurements – Electronic Compensation

Figure 9 indicates the problem encountered when propidium iodide (PI) and Fluorescein (FITC) are used in combination: FITC emission goes far into the yellow light, thus giving not only a FL-1 but also FL-3 signal. This creates the need for electronic compensation of the proportion of FITC signal from FL-3 channel, otherwise apoptotic cells would be inadvertently counted as dead cells since uncompensated signals from Annexin-V bound to exclusively apoptotic cells would also create a FL-3 signal.



**Figure 9 - Flow Cytometry Compensation:** A) Wavelength spectrum showing commonly used wavelengths and lasers for dye excitation. FITC is detected at 530nm (FL-1), PI is detected at 585nm (FL-3). B) The compensation problem: emission spectra of FITC and PI show a signal overlap for PI detection at 585nm, resulting in a wrong PI readout. In order to compensate this error FL-1 signal needs to be

deducted from FL-3 signal. C) Characteristic pot plots: the left plot shows PI stained populations. The lower left quadrant contains the viable population, while the upper left population shows dead cells, where PI has entered the cell and bound to DNA. The dot plot on the right shows the end-result of compensated Annexin-V/PI stain with an apoptotic population in the lower right quadrant (only Annexin is bound) and a dead population in the upper right (both PI and Annexin bind).

Briefly, compensation needs 4 extra samples and is done 4 steps using a logFL-1 / logFL-3 dot plot and via adjusting PMT settings for FL-1 and FL-3 and compensation settings FL3 -%FL-1:

Sample #1: unstained cells: adjust FL-1 and FL-3 between  $10^0$  and  $10^1$  so that the population is located in the lower left quadrant.

Sample #2: PI stain: adjust FL-1 PMT so that the viable population is located in the lower left quadrant and the dead population in the upper left quadrant. FL-1 signal should range between  $10^0$  and  $10^1$  similar to the left dot plot in Figure 9, section C.

Sample #3: Annexin-V-FITC stain: adjust FL-1 so that no cells are in the upper left or right quadrant (FL-3 between  $10^0$  and  $10^1$ ) by adjusting the FL3 - %FL1 compensation parameter.

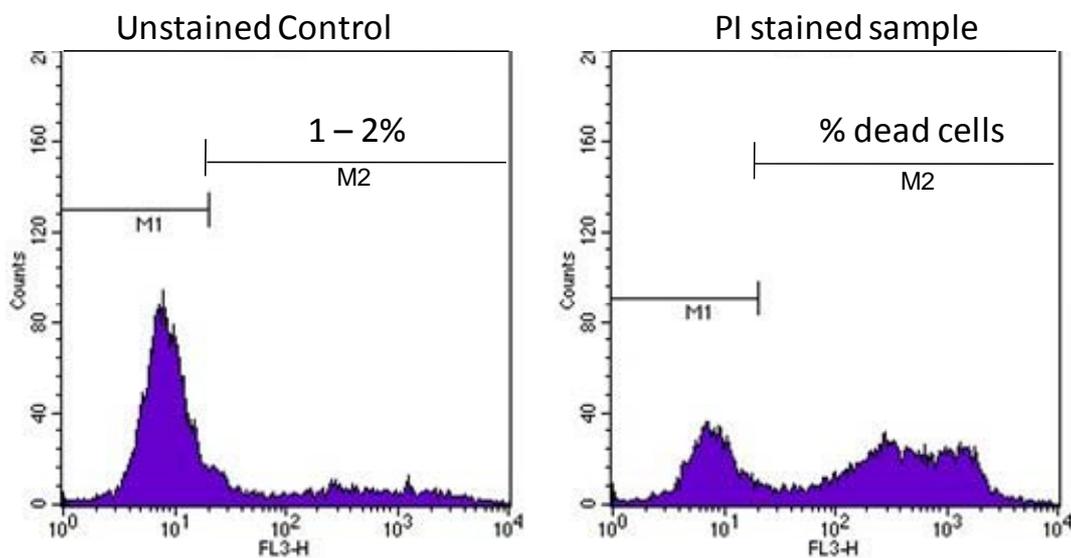
Sample #4: Annexin / PI stain: for control sample with apoptotic as well as dead cells. Correct compensation and adjustment of PMT values should give a dot plot similar to the right dot plot in Figure 9, section C.

#### 3.2.2.2. Viability – Propidium Iodide Stain

Propidium iodide (BD Biosciences, Pharmingen, San Diego, CA), an intercalating and fluorescent molecule, is a commonly used agent for DNA staining. PI is excited at a wavelength of 635nm and fluorescence can be detected at 585nm (within a range of 562nm to 588nm). Being most commonly used to quantitatively assess DNA content, PI is also useful for quick discrimination between viable and dead cells since leaky membranes of dead cell allow PI to enter the cell and bind DNA.

For flow cytometry analysis of viability cell samples 300µl to 500µL corresponding to an overall of  $1 \times 10^5$  to  $1 \times 10^6$  cells. PI stock (25 µg/ml in 1x PBS) was added to yield a final concentration of 0.25 µg/ml and cells were incubated for 5 minutes at room temperature before analysis. Unstained and untreated negative control samples were used to set the FL-1 and FL-3

intensities between 1 and 10 on the logarithmic scale. Dead cells showed up as a separate population on the FL-1 / FL-3 dot-plot, characterized by high FL-3 and low FL-1 intensities. For calculation of cell viability, cell counts were plotted over FL-3 intensities in a Histogram Plot (see Fig. 10) using BD Cellquest Pro Software (BD Biosciences, Pharmingen, San Diego, CA ). Using unstained negative control samples, markers were set to the dead cells were gated and expressed as percentage of total counts using BD FACSCalibur Software.



**Figure 10 – Data Analysis of propidium iodide (PI) detection of dead cells using Histogram Plots:** Unstained controls are used to set markers M1 and M2 so that PI positive cells make up 1-2% of total events (left Histogram). This percentage is the result of cell fragments that exhibit low fluorescence detected at FL-1. The same gates are then used on PI stained samples (right histogram) and yield information on the percentage of viable and dead cells.

Another approach used to for a quick estimate on cell viability was based on assessing cell shape, size and granularity using FSC and SSC parameters. The forward scatter / side scatter dot plot was adjusted using healthy cells so that the population of viable cells could be gated. For treated samples the same gate was then used to estimate the amount of viable cells.

### 3.2.2.3. Apoptosis – Annexin-V + Propidium Iodide Stain

In order to discriminate between live, dead and apoptotic cells, PI was used in combination with FITC labeled Annexin (Annexin-V-Fluos, Roche Diagnostics GmbH, Mannheim, Germany).

The calcium dependent protein Annexin-V has specific binding affinity to phosphatidyl-serine (PS), a membrane lipid that is located on the inner side of the asymmetric cell membrane. In the course of apoptosis, the cell goes through several morphological changes (as discussed on page x introduction), one of them being translocation of membrane lipids from inner to outer half. Thus, in the course of apoptosis PS gets presented on the outside of apoptotic cells where it is accessible to Annexin-V. Due to leaks in the cell membrane, Annexin-V can bind PS residues of dead cells as well, which is why PI is used in this assay to selectively stain dead cells.

For staining  $5 \times 10^5 - 1 \times 10^6$  cells were taken up and centrifuged at 170g for 10 minutes. Supernating media was discarded and cell pellets were re-suspended in 50 $\mu$ l of Annexin-V-FITC/PI staining solution (200ng/ $\mu$ l Annexin-V-FITC, 1000 ng/ $\mu$ l propidium iodide stock in Annexin-V-binding buffer) and incubated at room temperature for 10 minutes. Cells were then centrifuged again for 5 minutes at 170g, staining solution was discarded and cells were re-suspended in 300 $\mu$ l Annexin-binding buffer (see Table 1 for composition). Samples for compensation were prepared the same way, leaving out either Annexin-V-FITC or PI dye, or both according to the compensation strategy described above.

**Table 1 - Preparation of 250 ml Annexin-V Binding Buffer**

Composition	Preparation
10 mM HEPES/NaOH pH 7.4	595.5 mg HEPES
140 mM NaCl	2045.4 mg NaCl
5 mM CaCl <sub>2</sub>	183.775 mg CaCl <sub>2</sub> .2H <sub>2</sub> O

After dissolving the amounts in approximately 200 ml HQ water, the pH was adjusted to 7.4 using 25% NaOH; volume was then brought to 250 ml and buffer was sterile filtrated (0.22  $\mu$ m) and aliquoted to sterile 50 ml centrifuge tubes.

Samples were measured on BD FACS Calibur until a total of 10000 cells had been counted. Analysis was performed using CellQuest Pro software and quadrant stats analysis. Therefore markers were set dividing the FL-1 / FL-3 dot plot into four rectangular regions (quadrants). As can be seen in Figure 9 (section C), viable, apoptotic and necrotic populations come in 3 distinct quadrants. Quadrant statistics were used to calculate the percentage of each population relative to the total amount of 10000 events.

#### 3.2.2.4. MicroRNA Transfection Efficiency

Transfection efficiency of miRNA precursors and antagomirs was determined using FAM labeled negative control precursor-miRNAs or antagomirs (Ambion Inc., Austin, TX). Transfections were carried out according to the standard protocol described and cells were harvested 24 hours after transfection and centrifuged at 170g for 10 minutes. Cell pellets were washed twice in sterile 1x PBS to remove transfection complexes that had adhered to the cell surface and would bias the measurement. Cells were then re-suspended in 1x PBS and subjected to flow cytometry measurement: therefore PMT-values were adjusted using mock-transfected cells (that included transfection agent but lacked FAM labeled RNA) so that FL-1 intensities ranged between  $10^0$  and  $10^1$ . Positively transfected cells showed high FL-1 intensities as a result of FITC labeled miRNAs within the cells. Transfection efficiency was analyzed by gating FL-1 positive cells, and dividing this population count by the number of total counts. Viability was estimated using the FSC/SSC plot and untreated, healthy controls for gate adjustment.

#### 3.2.3. Metabolite and Amino Acid Analysis

Concentrations of glucose, lactate, glutamine and glutamate were measured using the YSI 7100 MBS system, a multiparameter bioanalytical system (YSI Life Sciences, Yellow Springs, OH). Detection of metabolites is based on enzyme electrode membranes, where enzymatic metabolite conversion is translated to electric currents, which can be measured and related to metabolite concentration.

For analysis cells were centrifuged at 170g for 10 minutes, and supernatant media was transferred to separate microcentrifuge tubes. Tubes were loaded into the sample spots on the YSI 7100 and a custom program to analyze glucose, lactate, glutamine and glutamate was started.

#### 3.2.4. EpoFc Product Titer Analysis - ELISA

Levels of secreted Epo-Fc fusion protein were measured using an in-house developed enzyme-linked immunosorbent assay (ELISA). Briefly, ELISA technology applies the principles of immunology and is widely used to quantitatively detect presence or absence of an antibody or an antigen in a solution. Many variants of ELISA assays exist for various diagnostic applications.

Most commonly special microtiter plates are pre-coated with an antibody specific for the protein of interest. When this protein is applied to the coated well, it gets selectively bound and is not lost during subsequent rounds of washing. The antigen is then detected by addition of a secondary antibody which recognizes another epitope of the antigen. This secondary antibody is usually conjugated with enzymes such as peroxidases, which convert certain substrates into detectable fluorophores. Using a standard curve the detected intensities can be related to the original concentration of antigen in the solution.

For Epo-Fc analysis, 96-well plates were pre-coated overnight using a goat-anti-human-gamma-chain antibody (I-7883 Sigma, St. Louis, MO). Cell culture supernatants were diluted between 1:10 and 1:100 and applied to the pre-coated plates. Epo-Fc titers were quantified in a  $2^8$  dilution series with a goat-anti-human-gamma-antibody conjugated with horseradish peroxidase (HRP, 62-8420 Zymed Laboratories, San Francisco, CA). Purified Epo-Fc was used as a standard.

### **3.3. MicroRNA-21 Overexpression and Inhibition Assays**

#### **3.3.1. Batch Cultivation**

Batch cultivation was carried out in 6-well format with cultivation times of 96h following miRNA transfection. Transfections were performed according to the standard transfection protocol. Besides pre-miR-21 and anti-miR-21 three controls were used: (i) a negative-control pre-miRNA (nc-pre-miR) with no target mRNA within the cell; (ii) mock transfected cells where transfection was mimicked by adding transfection agent without RNA to the cells; and (iii) untreated cells, meaning no transfection or miRNA agent had been added. For all time points, samples for total RNA extraction and Taqman qPCR analysis were taken in order to control miR-21 overexpression and knockdown. In regular intervals, usually every 24 hours, samples for analysis of cell density, viability, metabolite concentration and Epo-Fc product titer were taken. In addition the amount of apoptotic cells was determined during in the late phase of cultivation, at 72h and 96h. Using the equations previously described, specific growth rate, integral of viable cells, specific production rate and space-time-yield were calculated.

### 3.3.2. Stress Resistance Assays

#### 3.3.2.1. Camptothecin Treatment

The anti-cancer drug camptothecin (CPT) is a plant alkaloid successfully used for treating gastric, rectum and bladder tumors. Its action is based on inhibition of Topoisomerase I (Topo I) via formation of a ternary complex, which results in stabilization of an irreversible DNA/Topo I complex. Thus, CPT acts on rapidly dividing cells with high Topo I activity where it causes DNA damage eventually leading to cell death via apoptosis (Hinz, Helleday et al. 2003).

In this study CPT was used to induce DNA damage and cell death in CHO in order to establish a functional protocol for apoptosis detection using flow cytometry. In the next step cells that had earlier been transfected with pre-miR-21 or anti-miR-21 were treated with CPT to test anti-apoptotic effects of miR-21. Therefore, cells were transfected at day 0 with pre-miR-21, anti-miR-21 and nc-pre-miR and seeded at  $1 \times 10^5$  cells / ml; also, as additional controls, mock transfected cells and untreated cells were prepared and seeded at the same density. For a 10mM stock solution 3.48 mg CPT (Sigma Aldrich, St. Louis, MI; 348.4 g/mol) were dissolved in 1000  $\mu$ l pure DMSO to yield a 10 mM stock solution. Dilutions of 1:10, 1:100 and 1:1000 with DMSO yielded 1 mM, 100  $\mu$ M and 10  $\mu$ M stock solutions, respectively. Final concentrations used for CHO-EpoFc treatment generally ranged between 10 nM and 20  $\mu$ M and cells were treated according to volumes listed in Table 2.

**Table 2 – Camptothecin Treatment Scheme**

<b>Final CPT Concentration</b>	<b>10 nM</b>	<b>50 nM</b>	<b>100 nM</b>	<b>500 nM</b>	<b>1 <math>\mu</math>M</b>
<b>2500<math>\mu</math>l total volume</b>	1:1000 (10 $\mu$ M Stock) 2,5 $\mu$ L	1:200 (10 $\mu$ M Stock) 12,5 $\mu$ L	1:100 (10 $\mu$ M Stock) 25 $\mu$ L	1:200 (100 $\mu$ M Stock) 12, 5 $\mu$ L	1:100 (100 $\mu$ M Stock) 25 $\mu$ L
	<b>2 <math>\mu</math>M</b>	<b>5 <math>\mu</math>M</b>	<b>10 <math>\mu</math>M</b>	<b>15 <math>\mu</math>M</b>	<b>20 <math>\mu</math>M</b>
	1:500 (1 mM Stock) 5 $\mu$ L	1:200 (1 mM Stock) 12,5 $\mu$ L	1:100 (1 mM Stock) 25 $\mu$ L	1:666.67 (10 mM Stock) 3.75 $\mu$ L	1:500 (10 mM Stock) 5 $\mu$ L

CPT-treatment of transfected cells started 48h after transfection and viability and cell death were measured 8 hours after treatment. Approximately  $5 \times 10^5$  to  $1 \times 10^6$  cells/ml were taken up,

washed with 1x PBS and were stained with Annexin-V and propidium iodide as previously described.

### 3.3.2.2. Hydrogen Peroxide Treatment

Stress resistance of miRNA transfected cells was tested in response to addition of hydrogen peroxide causing oxidative stress within the cells. The influence of H<sub>2</sub>O<sub>2</sub> on an interferon- $\gamma$  producing CHO cell line has been assessed by Dunster et al. in 1997, with the result that hydrogen peroxide as well as other oxidative chemicals showed only little influence on cell growth as well as protein production (Dunster, Cheeseman et al. 1997). This observation was suggested to be the consequence of a long selection process of producer clones, making them more resistant. Yet, influence of H<sub>2</sub>O<sub>2</sub> on recombinant CHO-EpoFc in comparison to miRNA transfected CHO cells was tested in this study.

One percent as well as 0.1% hydrogen peroxide stock solutions were prepared from a 30% stock solution (Sigma-Aldrich, St. Louis, MI), corresponding to 294mM and 29.4mM stock solutions, respectively. Final concentrations in the range of 0.05 mM to 2 mM were prepared according to Table 3.

**Table 3– Hydrogen Peroxide Treatment Concentrations used in H<sub>2</sub>O<sub>2</sub> Treatment of CHO cells.**

<b>Final H<sub>2</sub>O<sub>2</sub> Concentration</b>	<b>0.05 mM</b>	<b>0.5 mM</b>	<b>1 mM</b>	<b>1.5 mM</b>	<b>2 mM</b>
<b>0.1% and 1% stocks 2500<math>\mu</math>L total volume</b>	1:588 4,25 $\mu$ L (0.1% stock)	1:588 4,25 $\mu$ L (1% stock)	1:294 8,5 $\mu$ L (1% stock)	1:196 12,75 $\mu$ L (1% stock)	1:147 17 $\mu$ L (1% stock)

For a preliminary experiments the impact of H<sub>2</sub>O<sub>2</sub> on CHO-EpoFc was assessed, cells were seeded at 3\*10<sup>5</sup> cells/ml and treated the same day with concentrations of H<sub>2</sub>O<sub>2</sub> according to Table 3. For transfection experiments, standard transfections including controls were performed; cells were seeded at 1\*10<sup>5</sup> cells/ml and incubated for 48h (where cell densities of approx. 3\*10<sup>5</sup> cells/ml had been reached) before treatment. Hydrogen peroxide treatment was performed as followed: respective amounts of H<sub>2</sub>O<sub>2</sub> (see Table 3) were added to each well and cells were incubated for 45 minutes in the presence of hydrogen peroxide. Cells were then centrifuged at 170g for 10 minutes, supernatant containing H<sub>2</sub>O<sub>2</sub> was discarded, and cell pellets

were re-suspended in 2.5 ml of fresh, pre-warmed media and seeded into new 6-well plates. After 8h incubation at 37°C cells were subjected to flow cytometry analysis for detection of cell viability, apoptosis and cell death.

### 3.3.2.3. Nutrient Depletion Assay

In recent studies the occurrence of apoptosis as well as autophagy at the end of batch cultivations, when glucose and glutamine becoming limiting substrates, was shown (Hwang and Lee 2008; Hwang and Lee 2008). In this study, the influence of miR-21 overexpression as well as knockdown on viability under glucose limiting conditions was measured.

CHO-EpoFc cells were cultivated in a glucose deprived DMEM/Ham's F12 media, containing only 1 g/L glucose instead of 3.4 g/L. After preliminary experiments testing the effect of glutamine concentrations in the range of 1 mM to 4 mM, it was decided on using 4 mM L-glutamine for the actual experiments. Cells were transfected with miR-21, anti-miR-21 and the corresponding controls, and seeded at  $1 \cdot 10^5$  cells/ml. Cell number and viability and metabolite concentrations were regularly monitored over a period of 144h, starting at 72h. Cultivations were stopped when viabilities had decreased below 50%.

## 3.4. RNA Analysis

### 3.4.1. Total RNA Isolation

Trizol reagent (Invitrogen, Carlsbad CA) is a mono-phasic solution of phenol and guanidine isothiocyanate that is used for homogenization or lysis of tissues or cells and subsequent RNA isolation. Trizol maintains the integrity of RNA while degrading and dissolving cells and cell components. Addition of chloroform followed by centrifugation leads to a separation in an aqueous and organic phase, with RNA being dissolved in the aqueous phase. The aqueous phase can be transferred to a fresh tube, and RNA can be precipitated by addition of isopropyl alcohol, which has the advantage that precipitation can be performed at room temperature.

All reagents were derived from Invitrogen (Carlsbad, CA) unless otherwise stated. Total RNA from cells was isolated using Trizol. For attached cell lines, cells were washed once with sterile 1x PBS to remove residual media and serum traces, and 1ml Trizol was added per 10 cm<sup>2</sup> of

covered surface area. Cells were lysed for 2 minutes in the flask until no attachment was visible. The lysate was transferred to a nuclease free 50 ml centrifuge tube, mixed vigorously and frozen at -80°C. For suspension cells, aliquots of 5 million cells were centrifuged at 170xg for 10 minutes. Supernatant was removed and cell pellets were re-suspended in 1 mL Trizol, mixed vigorously for 15 seconds on a vortex and incubated at room temperature for 5 minutes. Cell lysates were then briefly but vigorously mixed and stored at -80°C.

For total RNA isolation, Trizol cell lysates were thawed and 0.2 ml of chloroform was added to 1ml Trizol, mixed on a vortex for 15 seconds and incubated at room temperature for 3 minutes. Samples were then centrifuged at 12000 g for 15 minutes at 4°C to separate precipitated DNA and proteins from RNA. The upper aqueous phase which predominately contained dissolved RNA was transferred to a new nuclease free tube and extraction was repeated by adding 0.5 ml of Trizol followed by 0.2 ml of Chloroform (double extraction for increased RNA purity/quality). The samples were mixed for 15 seconds on a vortex, incubated at room temperature for 3 minutes and centrifuged again at 4°C and 12000 g for 15 minutes. The aqueous upper phase was transferred to a new nuclease free tube and 0.5 ml 100% isopropanol (0.22µm filtered) was added per 1 ml Trizol. Samples were mixed well and incubated at room temperature for 10 minutes, followed by centrifugation at 4°C and 12000g for 10 minutes. Isopropanol was then discarded and RNA pellets were washed using 1ml 70% Ethanol (0.22µm filtered) and centrifuged for 5 minutes at 7500 g at 4°C. Ethanol was then carefully removed and pellets were dried for about 10 – 15 minutes in a laminar flow hood. RNA was re-suspended in 20 – 30 µl nuclease free water and was stored at -80°C.

#### **3.4.2. RNA Quality Assessment**

Agilent 2100 Bioanalyzer in combination with Agilent RNA 6000 Nano kit was used in order to assess to quality of isolated total RNA. This “Lab-on-a-chip” system applies microfluidics, capillary electrophoresis and fluorescent dyes in order to evaluate concentration and integrity of nucleic acids. RNA molecules that are applied to sample wells on the chip move through a separation channel where they get bound by intercalating dyes. The fluorescence is then measured when these molecules pass a detector, and is proportional to mass via the peak size,

and size via retention time. Since ribosomal RNA (rRNA) makes up more than 80% of total RNA in eukaryotes, mainly 18S and 28S rRNA, it is commonly assumed that rRNA quality reflects that of mRNA as well. In general a ratio of 28S:18S rRNA of 2:1 is considered to represent high quality RNA, as well as the overall peak distribution of total RNA profiles.

For RNA quality analysis using Agilent RNA 6000 Nano Chips, 2µL aliquots were taken from total RNA samples and diluted with nuclease-free water to a concentration in the range of 5 ng/µl and 500 ng/µl. Total RNA samples were then incubated at 70°C for 2 minutes in order to denature any RNA secondary structures. 1 µl of denatured total RNA was applied to each sample well of a RNA 6000 Nano Chip which had previously been prepared for analysis according to the manufacturer's protocol. RNA integrity was evaluated based on peak size and shape as well as 28S:18S peak ratios in a time vs. fluorescence chromatogram.

### **3.4.3. Total RNA Quantification**

Concentration of RNA, DNA and proteins are commonly analyzed using the characteristic absorption profiles of nucleic acids and proteins. RNA and DNA are generally detected at wavelengths of 260 nm, whereas proteins as well as phenols are detected at 280 nm. Photometrical measured absorbance values can then be related to concentration using Lambert-Beer's law:

$$A = \epsilon * l * c$$

*A ... absorbance,  $\epsilon$  ... extinction coefficient, l ... path length, c ... concentration*

Thus, for quantification of RNA in a sample A<sub>260</sub> is used, whereas A<sub>260</sub>/A<sub>280</sub> gives an idea of the purity of isolated total RNA, and A<sub>260</sub>/A<sub>280</sub> values should at least exceed values of 1.7 for qPCR analysis and 1.8 for RNA Chip analyses.

At the University of Minnesota, concentration of total RNA was measured using a SpectraMax Plus (Molecular Devices, United States) 384-well plate reader. Total RNA samples were diluted 1:40 and 1:80 with nuclease free water, and 40µL were applied to the wells of 384-well plate.

OD-values between 0.1 and 1.0 at 260nm and 280nm were considered for concentration calculation based on Lambert-Beer's Law using an extinction coefficient of 38.1.

At Boku University in Vienna, total RNA concentration was determined using IMPLEN NanoPhotometer 7122 v1.6.1 (Implen GmbH, Munich, Germany). For this analysis the LableGuard Application for RNA analysis (LID Factor 10, Calculation factor 40) was used. All measurements were background corrected using a nuclease-free H<sub>2</sub>O blank. Samples were diluted to match a concentration range of 20 to 800 ng/μl total RNA, and 3 – 5 μl of sample were applied to the cuvette.

OD at 260nm was used for RNA concentration calculation, and the OD 260/280 ratio was taken to estimate purity from proteins and phenol as described before.

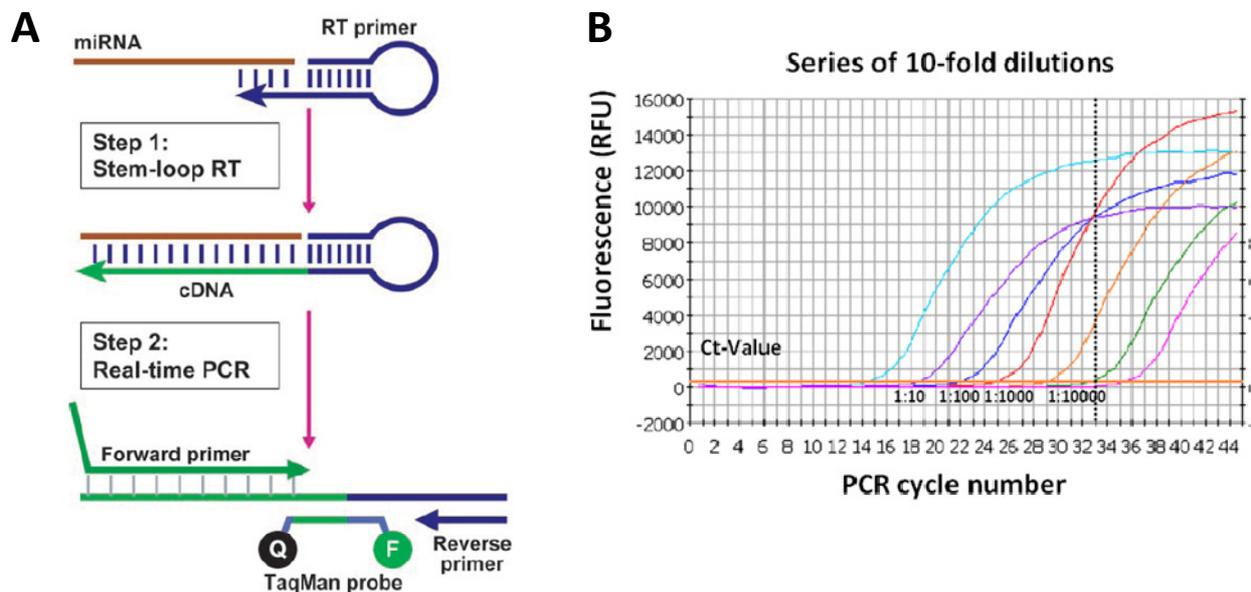
#### **3.4.4. Taqman quantitative real time PCR**

Quantitative real-time polymerase chain reaction (qPCR) methods are based on the common PCR amplification technique combined with photometric assessment of each amplification round. Briefly, traditional PCR allows exponential amplification of DNA molecules using temperature-stable DNA polymerases, such as Taq DNA polymerase derived from *Thermus aquaticus*. For an amplification reaction, the thermo-stable polymerase needs to be combined with short oligo-nucleotides (approximately 20 nt long, referred to as primers), that are complementary to defined DNA sequences on each of the two strands of the DNA of interest, and with the DNA template. The amplification reaction is controlled by using specific temperatures: one amplification cycle consists of denaturation (95°C), annealing (approx. 50-60°C depending on specificity) and extension (approx. 72°C). Repeating this cycle, results in logarithmic amplification of the DNA sequence of interest. After about 40 cycles most reactions are stopped, and the amplified DNA is applied to an agarose gel for analysis and purification. PCR reactions like this are at the best semi-quantitative, and cannot be used to relate the amount of PCR product to initial amount of PCR template. This is partly because ethidium bromide (EtBr), which is used for DNA staining in agarose gels, is an insensitive stain, and its intensity cannot be precisely related to amount of DNA; also, after 30 – 40 amplification cycles

exponential amplification has already ended, meaning the final amount of amplified DNA does not directly correspond to initial amounts as can be seen in Figure 11, section B.

Quantitative PCR or real-time PCR (TaqMan PCR) technology applied in this study uses additional probes that are linked with FRET (fluorescence resonance energy transfer) dyes; FRET refers to a donor-fluorochrome (reporter) transferring energy to a close receptor molecule (quencher) when it is excited by a light source. Only if the distance between reporter and quencher is increased, light that is emitted from the reporter is not transferred to the quencher and can therefore be detected. In case of TaqMan real-time PCR, Taq polymerases with 5'-3' exonuclease activity are used which, during polymerization, remove the specific FRET-probe bound to mature miR-21. This removal results in increased distance between reporter and quencher and light emitted from the reporter can be detected. Consequently, in the course of a PCR reaction, the amount of light emitted increases with every cycle. For quantitative analysis only the exponential amplification phase (straight line in the cycle/log(cDNA) plot) is of interest. The specificity of Taqman PCR relies on the very specific reverse transcription of only mature miRNAs into cDNA. This is achieved by using TaqMan® MicroRNA Reverse Transcription Kit, where a stem-loop primer that selectively binds only to the mature miRNA of interest is used; in the course of reverse transcription, this stem-loop primer is extended to yield a mature miRNA cDNA plus loop (Chen, Ridzon et al. 2005). The principles of Taqman PCR are summarized in section A of Figure 11.

After running qPCR on several samples the outcome is a graph that depicts fluorescence intensities for each sample over PCR cycle number. Figure 11, B gives an example for a series of 10-fold dilutions of the same sample. In the next step the user sets a specific level of fluorescence which is significantly higher than the background fluorescence as threshold; the PCR cycle at which this threshold level of fluorescence is crossed is called *cycle threshold value*, or *Ct-value*. This value directly corresponds to the initial amounts of cDNA, and a difference of 1 in Ct-values means double or half amount of initial cDNA.



**Figure 11 – (A) Schematic description of TaqMan miRNA assays:** TaqMan-based real-time quantification of miRNAs includes two steps, stem–loop RT and realtime PCR. Stem–loop RT primers bind to the 3' portion of exclusively mature miRNA molecules and are reverse transcribed with reverse transcriptase. Then, the RT product is quantified using conventional TaqMan PCR that includes miRNA-specific forward primer, reverse primer and dye-labeled TaqMan probes. The purpose of tailed forward primer at the 5' end is to increase its melting temperature ( $T_m$ ) depending on the sequence composition of miRNA molecules. Figure taken from (Chen, Ridzon et al. 2005). (B) This figure shows qPCR runs of ten-fold dilutions of a cDNA sample, resulting in lines that cut the Ct-value at a regular interval, following the exponential trend. The Ct-value is set at the lowest fluorescence value significantly different from background fluorescence. At the end of the PCR reaction (cycles 30 – 40) initial amounts of cDNA do not correspond to the fluorescence signal anymore, showing that common PCR reactions are at the best semi-quantitative.

In order to compensate for different starting concentrations certain references have to be co-measured; these references are often house-keeping genes that need to fulfill the following requirements: (i) all cells need to express this house-keeping gene, (ii) all cells need to possess the same copy number and (iii) the expression level of this gene needs to remain constant under various conditions. In this study U6 smallRNA was used as a reference. Thus, each RNA sample was reverse transcribed for U6 small RNA as well as for mature miR-21. After qPCR detection runs for both U6 and miR-21, delta Ct-values were calculated by subtracting the U6 Ct-value from the corresponding miR-21 Ct-value of the same sample. In order to compare different miRNA transfections the miR-21 induction for each transfection was related to the

miR-21 level in cells transfected with a non-coding control; this was done via calculation of a deltadelta Ct-value. All calculations are summarized below for the case of pre-miR-21 and anti-miR-21 transfected cells.

$$\text{deltaCt}_{miR21} = Ct(miR21) - Ct(U6) \dots \text{from pre\_miR\_21 total RNA}$$

$$\text{deltaCt}_{anti\_miR21} = Ct(miR21) - Ct(U6) \dots \text{from anti\_miR\_21 total RNA}$$

$$\text{deltaCt}_{nc\_miR21} = Ct(miR21) - Ct(U6) \dots \text{from nc\_pre\_miR total RNA}$$

$$\text{deltadeltaCt}_{miR21} = \text{deltaCt}_{miR21} - \text{deltaCt}_{nc\_miR21}$$

$$\text{deltadeltaCt}_{anti\_miR21} = \text{deltaCt}_{anti\_miR21} - \text{deltaCt}_{nc\_miR21}$$

$$\text{deltadeltaCt}_{nc\_miR21} = \text{deltaCt}_{nc\_miR21} - \text{deltaCt}_{nc\_miR21}$$

A deltadeltaCt-value of +/- 1 means half or doubled amount of cDNA, while a value of +/- 2 means fourfold decrease or increase in cDNA respective to the control. Thus fold change is calculated using the exponential equation  $2^{(\text{deltadeltaCt})}$ , as described in Table 4 below.

**Table 4 – Fold change calculation based on deltadeltaCt-values**

Taqman miRNA qPCR	nc-pre-miR	pre-miR-21	anti-miR-21
<b>Equation</b>	$2^{(-\text{deltadeltaCt}_{nc\_miR21})}$	$2^{(-\text{deltadeltaCt}_{miR21})}$	$2^{(-\text{deltadeltaCt}_{anti\_miR21})}$
<b>Fold change (sample/control)</b>	1	> 1 <i>(theoretically)</i>	< 1 <i>(theoretically)</i>

Total RNA samples from Trizol extraction were diluted to 10 ng/μl in nuclease-free water (Ambion Inc., Austin, TX). In the next step mature miR-21 and U6 small RNA were selectively reverse transcribed into cDNA in each sample, using Ambion looped reverse transcription primers. Thus, for each sample two mastermixes (compositions see Table 5) were prepared, one using miR-21 primers, and one using U6 primers:

**Table 5 - Reverse Transcription mastermix**

Mastermix für Reverse Transcription miR-21	Volumes for a single reaction ( $\mu$ l)
RNase free water	8.16
10x Reverse Transcription Buffer	1.5
dNTP Mix (100 mM)	0.15
RNase Inhibitor (20 U/ $\mu$ l)	0.19
5x TaqMan RT Primer (either U6 or miR-21)	3
Multiscribe Enzyme (50 U/ $\mu$ l)	1
template RNA (10ng/ $\mu$ l)	1
Total volume	15

Mix was kept on ice all the time and was centrifuged for 10 seconds prior to starting the reverse transcription reaction:

30 minutes  $\rightarrow$  16°C

30 minutes  $\rightarrow$  42°C

5 minutes  $\rightarrow$  85°C

After denaturation step, samples were put immediately on ice, ore frozen at -20°C. For Taqman real-time PCR each sample was run in quadruplicate and U6 and miR-21 mastermixes were prepared for each sample as described in Table 6.

**Table 6 - Real-time PCR mastermix**

Mastermix für Taqman real-time PCR miR-21	Volumes for a single reaction ( $\mu$ l)
20x TaqMan microRNA Assay qPCR Primer	0.5
TaqMan 2x Universal PCR Master Mix No AmpErase UNG	5
RNase free water	3.17
Product from RT reaction	1.33
Total volume	10

After preparing mastermixes, 8.67 $\mu$ l were aliquoted to Corbett qPCR tubes (former Corbett Life Sciences now Qiagen, Germay) on a 4°C cooled tube rack. Finally 1.33  $\mu$ l of cDNA were added. 64 samples could be analyzed at a time, and for each run non template controls were included,

where 1.33 µl nuclease-free water were added instead of cDNA. qPCR tubes were then inserted into a 72-tube rotor and put on the Corbett Rotor Gene 6000 qPCR instrument (now Qiagen Rotor-Gene-Q, Qiagen, Germany). Data analysis was done by applying  $\Delta\Delta Ct$  normalization as previously described (3.4.4).

### **3.5. MicroRNA MicroArray Profiling**

#### **3.5.1. Exiqon LNA MicroArrays (BOKU University)**

The difficulty with microRNA MicroArray technology is the shortness of microRNAs as target molecules as well as the great variance in base composition of miRNAs (i.e. the GC content can range from 25% to 90%). As a consequence the requirement for probes on the array to have similar melting temperature ( $T_m$ ) is difficult to meet since this would result in short probes (8-9 nt) for GC rich duplexes which again results in reduced specificity. In order to get good levels of affinity and specificity for the same melting temperature across all probes, Exiqon introduced a new technology called locked nucleic acid (LNA). The term locked nucleic acids refers to a synthetic RNA/DNA analogue where the ribose ring is locked by a methylene bridge connecting the 2'-O atom and the 4'-C atom. This bridge locks the ribose in the 3'-endo structural conformation and increases the thermostability of the duplex. By varying the LNA content it is possible to make probe lengths and  $T_m$  values more uniform which allows hybridization conditions that suit all miRNAs. It also enables efficient discrimination between closely related miRNAs and this detection sensitivity also eliminates the need for small RNA enrichment, as the miRNAs are 3' end labeled in the total RNA sample.

##### **3.5.1.1. Array Design**

The capture probes were spotted by Marcel Scheideler from the Technical University in Graz. Probes, based on Sanger miRBase version 9.2 (May 2007) were printed on epoxy-coated slides and include 394 known human miRNAs, U6 snRNA capture probes (again endogenous gene controls for signal normalization), negative control probes and Hy3 pre-labeled ankerspots to simplify GenePix Array List File (GAL-File) orientation.

### 3.5.1.2. MicroRNA Labeling

Labeling of miRNAs for differential expression analysis on miRNA Arrays was performed according to Exiqon miRCURY™ LNA microRNA array labeling kit. Molecules were labeled with a single fluorophore per molecule by applying a 2-step procedure directly on isolated total RNAs: first Calf Intestinal Alkaline Phosphatase (CIP) was used to remove 5' phosphates from RNA molecules, followed by enzymatical attachment of the fluorescent label to the 3'-end of miRNAs and enzyme inactivation.

**A**

RNA sample #	Sample description	Sample concentration (ng/μL)	μL RNA	μL H2O	Total
1	3F8-37°C	819	1.22	0.78	2μL (=1μg RNA)
2	3F8-33°C	656	1.52	0.48	2μL (=1μg RNA)
3	dhfr-37°C	810	1.23	0.77	2μL (=1μg RNA)
4	dhfr-33°C	799	1.25	0.75	2μL (=1μg RNA)
5	K1PD-37°C	673	1.49	0.51	2μL (=1μg RNA)
6	K1PD-33°C	559	1.79	0.21	2μL (=1μg RNA)
7	14F2-37°C	558	1.79	0.21	2μL (=1μg RNA)
8	14F2-33°C	344	2.91	-0.91	2μL (=1μg RNA)

**B**

CIP RXN Mastermix	Volume (μL)
total RNA (not in MM)	2
H <sub>2</sub> O (instead of spike-in)	1
CIP Buffer	0.5
CIP Enzyme	0.5
<b>Total Volume</b>	<b>4</b>

**C**

Mastermix Cy-5	Volume (μL)	Mastermix Cy-3	Volume(μL)
CIP Reaction	4.0	CIP Reaction	4.0
Labeling buffer	3.0	Labeling buffer	3.0
<b>dye Cy-5</b>	1.5	<b>dye Cy-3</b>	1.5
DMSO	2.0	DMSO	2.0
Labeling enzyme	2.0	Labeling enzyme	2.0
<b>total</b>	<b>12.5</b>	<b>total</b>	<b>12.5</b>

**Figure 12 – MicroArray samples and Mastermixes:** (A) overview of samples and sample concentrations. Each RNA was diluted to 0.5 μg/μl. (B) Calf intestinal alkaline phosphatase (CIP 10 U/μl) reaction mix for removal of 5'-phosphate residues. Amount of total RNA should be in the range of 0.25 to 1μg. (C) Cy-5 (37°C samples 1,3,5 and 7) and Cy-3 (33°C samples 2,4,6 and 8) labeling reaction mixes.

The lyophilized labeling dyes Hy3 and Hy5 were dissolved in 29 μl nuclease-free water, mixed on a vortex and centrifuged briefly to collect the full volume. Throughout the labeling procedure light exposure of dyes was minimized in order to avoid photodegradation of dyes.

Also dye damage can be the result of oxidation due to elevated ozone levels in the atmosphere, therefore demanding quick and careful handling of all reagents. All kit components were thawed on ice for 15-20 minutes; total RNA was diluted to 0.5 µg/µl and for the CIP reaction was prepared according to the manufacturer's protocol (see Figure 12, B).

The reaction mix was incubated at 37°C for 30 minutes followed by 5 minutes at 95°C and immediately transferred on ice (for up to 15 minutes). Using 4 µl from the CIP reaction the labeling reaction was prepared (see Figure 12, C), gently mixed on a vortex and briefly centrifuged. Samples were incubated in the absence of light at 16°C for 60 minutes followed by enzyme inactivation at 65°C for 15 minutes; labeled small RNA samples were then put to 4°C until further use. Hy 3 and Hy 5 labeled samples (1/2; 3/4; 5/6; 7/8 according to Figure 12, A) were pooled (25 µl) for this two dye experiment. Then 25 µl nuclease-free water and 50 µl Exiqon 2x Hybridization Buffer were added.

### 3.5.1.3. Hybridization

Hybridization was performed according to the microRNA Chip Hybridization manual for Tecan HS400 instrument. Therefore, the buffers listed in Table 8 had to be prepared and filtered using 0.22µm bottle-top filters (Millipore). All channels were primed following the manual and microRNA LNA chips were inserted. Filtered Pre-Hybridization Buffer was pre-heated to 64°C for 10 minutes before reverse pipetted onto the slides. Before applying 75 µl of labeled sample to the arrays, the samples were denatured for 3 minutes at 90°C, centrifuged and injected onto the arrays in the hybridization station. Hybridization was performed at 64°C for 16 hours. Table 7 gives a summary of samples and slide barcodes used for this experiment.

**Table 7 – MicroArray Barcodes and samples**

Array Barcodes	Cy3 (532 nm)	Hy 5 (635 nm)
03020379	3F8-33°C	3F8-37°C
03020379	dhfr-33°C	dhfr- 37°C
03020379	K1PD-33°C	K1PD-37°C
03020379	14F2-33°C	14F2-37°C

**Table 8 - Wash and Hybridization Buffers for Exiqon LNA MicroArrays**

<b>Prehybridization Buffer, 50 ml</b>			
Volume Percentage	Volume	Starting Material	Final Concentration
25%	12,5 ml	20x SSC (Invitrogen)	5x SSC
1%	0,5 ml	10% SDS (Invitrogen)	0,5 % SDS
	0,5 g	BSA (ROTH)	1% BSA
74%	37 ml	MilliQ Water	

<b>Wash Buffer A (2x SSC, 0,2% SDS) connected to liquid channel 2, 500 ml</b>			
Volume Percentage	Volume	Starting Material	Final Concentration
10%	50 ml	20x SSC (Invitrogen)	2x SSC
2%	10 ml	10% SDS (Invitrogen)	0,2% SDS
88%	440 ml	MilliQ Water	

<b>Wash Buffer B (1x SSC) connected to liquid channel 5, 500 ml</b>			
Volume Percentage	Volume	Starting Material	Final Concentration
5%	25 ml	20x SSC (Invitrogen)	1x SSC
95%	475 ml	MilliQ Water	

<b>Wash Buffer C (0,2x SSC) connected to liquid channel 6, 500 ml</b>			
Volume Percentage	Volume	Starting Material	Final Concentration
1%	5 ml	20x SSC (Invitrogen)	0,2x SSC
99%	495 ml	MilliQ Water	

<b>0,2% SDS connected to liquid channel 1, 500ml</b>			
Volume Percentage	Volume	Starting Material	Final Concentration
2%	10 ml	10% SDS (Invitrogen)	0,2% SDS
98%	490 ml	MilliQ Water	

#### 3.5.1.4. Scanning

After slide drying, the arrays were scanned on a Genepix 4000B array scanner (Axon Instruments, Union City, CA). The scanner uses two lasers to excite the fluorophores Hy3 (532nm) and Hy5 (635nm) and a pair of high-sensitivity photomultiplier tubes (PMTs) to detect spot fluorescence. The arrays were scanned with 5  $\mu$ m resolution and 100% laser power for both channels. PMT settings were adjusted that the majority of the features fall between 15.000 and 50.000 intensity units in the pixel intensity histogram. The Hy5 and the Hy3 channels were balanced so that the count ratio is close to 1. Oversaturation (intensity units above 65.000) was avoided. PMT voltages used were 700 for Hy3 and 800 for Hy5.

#### 3.5.1.5. Data Analysis

Intensity values from each feature were extracted using GenePixPro 4.1. software (Axon Instruments, Union City, CA) and a Gene Pix Array List file (GAL-file) which specifies the names, IDs and position of each spotted probe. Features containing obvious artefacts were tagged with a flag and not included in further analysis. The intensities for all intact spots were saved in a Gene Pix Result file (GPR-file) without further data processing, since only results for Cy3 were further analyzed (see results section for rationale). Using Microsoft Excel, background corrected median feature intensities were extracted for each spot. “Empty” feature, Hy3 control spots as well as flagged spots were filtered from the dataset, and the remaining features were analyzed for intensity characteristics (mean, median and quartile values). The Pivot Table function was used for calculating a mean value and standard deviation for each ID (spotted probe). Subsequently for a presence/absence analysis, all features with a standard deviation greater 20% were filtered in order to guarantee a high quality of intensity values. The remaining miRNAs (or control small RNAs) and mean intensity values were extracted and counted. For each library the intensity of high quality spots was normalized to U6-snRNA intensity. MicroRNA expression as well as absence/presence analysis was performed in Microsoft Excel/Visual Basic or the open-source statistic tool *R*.

### 3.5.2. Ambion mirVANA MicroArrays (University of Minnesota)

#### 3.5.2.1. Small RNA Isolation for Microarray Analysis

100µg of total RNA were diluted in 100µL TRIS-HCl pH 7.8 (0.22µm filtered) and denatured for 5 minutes at 70°C. Samples were then applied to a YM-100 microcon column (Millipore) and centrifuged at room temperature for 30 minutes at 5000rpm or until no more liquid passed through the column. The eluate containing the small RNAs up to 300nt in length was recovered. 0.1 volumes of 5M NaCl (0.22µm filtered) were added to adjust the salt concentration to 0.5M. Linear Acrylamide (Ambion Inc., Austin TX) was added to obtain a final concentration of 10 – 20 µg/ml and finally 2.5 volumes of ice-cold 100% Ethanol were added. After mixing thoroughly small RNAs were precipitated at -20°C for at least 3 hours or up to 16 hours (overnight).

Samples were centrifuged at 12000g and 4°C for 30 minutes, pellets were washed with 0.5ml 75% Ethanol (0.22µm filtered) and dried up side down in a laminar flow hood. After 10 – 15 minutes pellets were resuspended in 20µl nuclease free water.

#### 3.5.2.2. MicroRNA labeling

MicroRNAs were labeled using the MirusBio Label IT Kit miRNA labeling kit (MirusBio, Madison, WI). Small RNA isolated from 200µg total RNA was used for a single labeling reaction. Therefore the volume of re-suspended small RNA was brought to 86µl using nuclease free H<sub>2</sub>O. 10µl of 10x Labeling Buffer M were added followed by 2µL of Label IT Cy3 or Cy5 reagent (MirusBio, Madison, WI). From this time on samples were kept in the dark, and were handled in the absence of artificial light, in order to avoid dye degradation.

Labeling mixes were incubated at 37°C for 60 minutes and the labeling reaction was stopped by adding 0.1 volumes (10µl) of 10x Stop Solution. Labeled microRNAs were precipitated by adding 2µl of precipitation enhancer solution (MirusBio, Madison, WI), 0.1 volumes of 5M NaCl and 2.5 volumes of ice-cold 100% Ethanol. The mixture was incubated at -20°C for 3h and centrifuged at 12000 g and 4°C for 30 minutes. Pellets were washed using 75% Ethanol and dried upside down in a sterile laminar flow hood. Labeled miRNAs were then re-suspended using pre-warmed 1x Hybridization Buffer (MirusBio, Madison, WI).

#### 3.5.2.3. MicroRNA MicroArrays

Slides were spotted at BTI Singapore with Ambion mirVANA probe set V1 which is based on Sanger miRBase release 8.0 (October 2005). A total of 384 probes were spotted on the arrays, varying between 42 – 46 nucleotides in length. Each probe contains 18 – 24nt long specific segment that target miRNAs from either human, mouse or rat. In total 330 human (281 conserved, 49 predicted), 49 mouse and 14 rat miRNAs were represented on one slide. as well as 7 control spots.

#### 3.5.2.4. Slide Post Processing

First the spotted area was marked at the back of the slide, then slides were loaded onto a hydration chamber with the spotted area facing down and hydrated for 30 seconds to 1 minute

at 42°C using 0.5x SSC. Slides were then washed at room temperature to remove unbound probes using 0.1% Triton X-100 (5 min), 1mM HCl (2 min), 100mM KCl (10 min) and were rinsed for 1 min using de-ionized water (diH<sub>2</sub>O). Slides were immediately subjected to blocking at 50°C for 15 minutes. For blocking a 0.1M Tris HCl buffer with 100mM Ethanolamine and 0.1% SDS, which had been prepared freshly, was used. Slides were then rinsed with diH<sub>2</sub>O and dried using a centrifuge at 700rpm for 15 minutes. Dried slides were stored in a desiccator until used for hybridization.

#### 3.5.2.5. Hybridization

16µl of labeled miRNA were mixed 16µl 2x hybridization buffer (50% de-ionized formamide, 10x SSC, 0.2% SDS, 0.22µm filtered), heated to 95°C for 3 minutes and cooled to room temperature. Then Cy3 and Cy5 labeled miRNAs were pooled to give a total volume of 60µl. Lifter slips (22x50, Erie Scientific) were cleaned using 75% Ethanol and freed from dust, cautiously put onto a post-processed slide, so that the protruding edges faced down. Labeled miRNAs were then applied to an open end of the lifter slip, so that the surface covered by the lifter slip was completely covered with hybridization mixture. Slides were then sealed in a hybridization chamber and put to 45°C in a water bath for 16h. 10µl of 3x SSC were pipette into spots within the hybridization chamber in order to avoid dehydration during the 16 hours of hybridization. Slides were taken out of the hybridization chamber and submerged in 1x SSC / 0.2% SDS to remove the cover slip. Slides were washed for 10 minutes in 2x SSC, followed by 10 minutes 0.1x SSC / 0.1% SDS and 2 minutes 0.1x SSC and finally dried during centrifugation at 700rpm for 15 min.

#### 3.5.2.6. Scanning

After slide drying, the arrays were scanned on a Scan Array 5000 (Perkin Elmer, Wellesley, MA). Cy3 (532 nm) and Cy5 (635 nm) dyes were excited, and fluorescence was detected using 10 µm scanning resolution and 100% laser power for both channels.

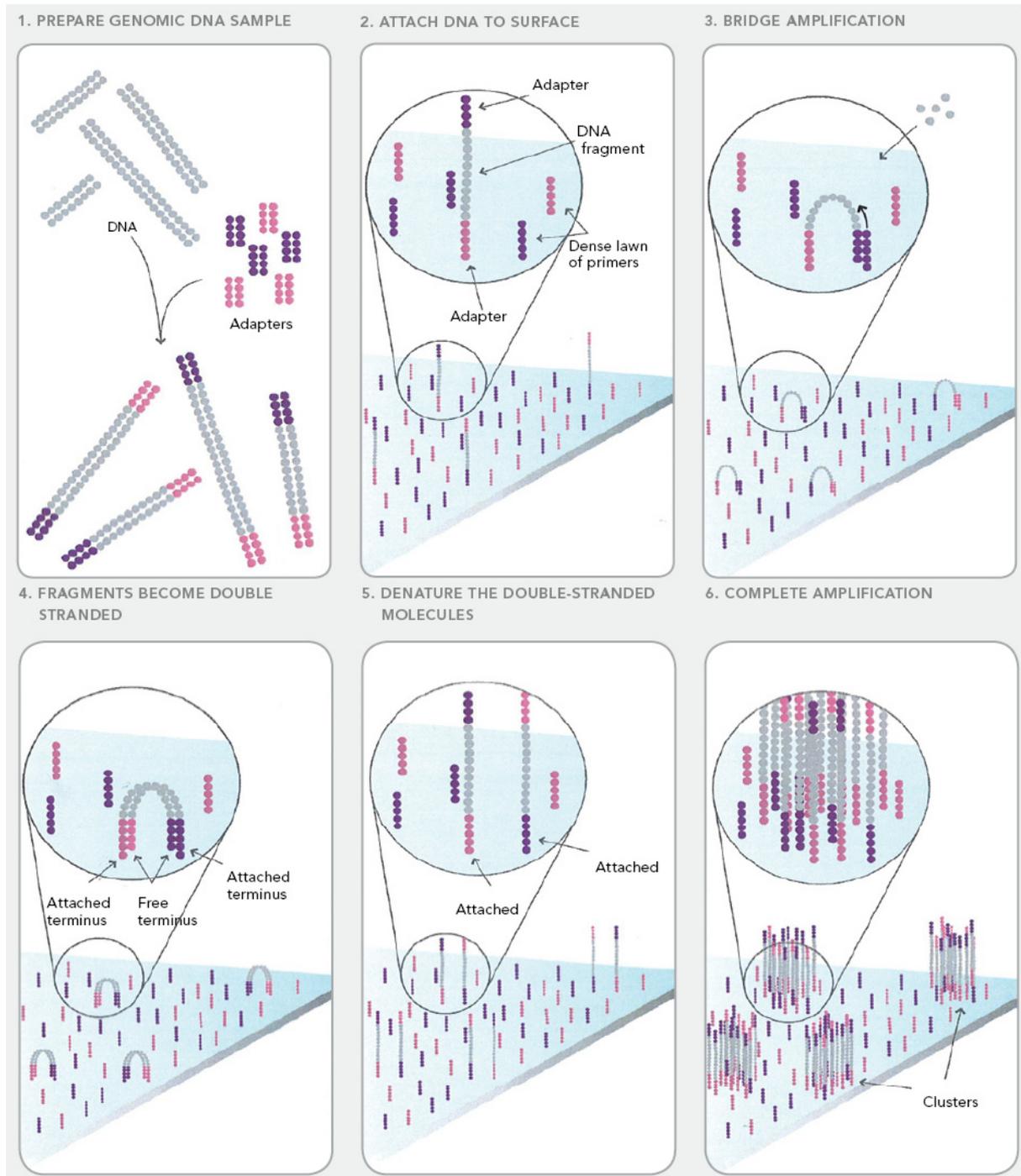
## 3.6. Illumina/Solexa Sequencing

### 3.6.1. Background

Illumina/Solexa sequencing technology is a cloning-free approach to large scale DNA sequencing. Figure 13 was taken from the Solexa/Illumina handbook and visualizes the crucial steps of this massive parallel sequencing technology:

In case of small RNA sequencing isolated RNA molecules are ligated to specific 3' and 5' adapters, reverse transcribed to cDNA and enriched using PCR. This results in double-stranded, blunt-ended, molecules with distinct adaptor sequence on either end. These double-stranded cDNA molecules are then applied to a single-molecule array (referred to as *flow cell*) where they become denatured and attached to the surface via their 3' or 5' adapters. The single-stranded attached molecules bend over and hybridize to complementary adaptors creating a bridge ("bridge-amplification") that functions as a template for synthesis of the complementary strand. Multiple rounds of annealing, extension and denaturation result in clonal clusters of about 1  $\mu\text{m}$  in diameter. About 40 million clusters per flow cell are generated with each cluster consisting of about 1000 copies of a single template molecule. DNA in each cluster is then linearized and denatured generating single-stranded templates for sequencing by synthesis: four reversible terminators (A, C, G and T), each labeled with a different fluorescent dye, together with primers and polymerases are added to the flow cell. Only the complementary base-terminator gets incorporated and gives a signal after laser excitation (Bentley, Balasubramanian et al. 2008; Morozova and Marra 2008). At the moment this technology can generate up to 36-bp reads using the 1G genome analyzer from Illumina (Illumina Inc., San Diego, CA).

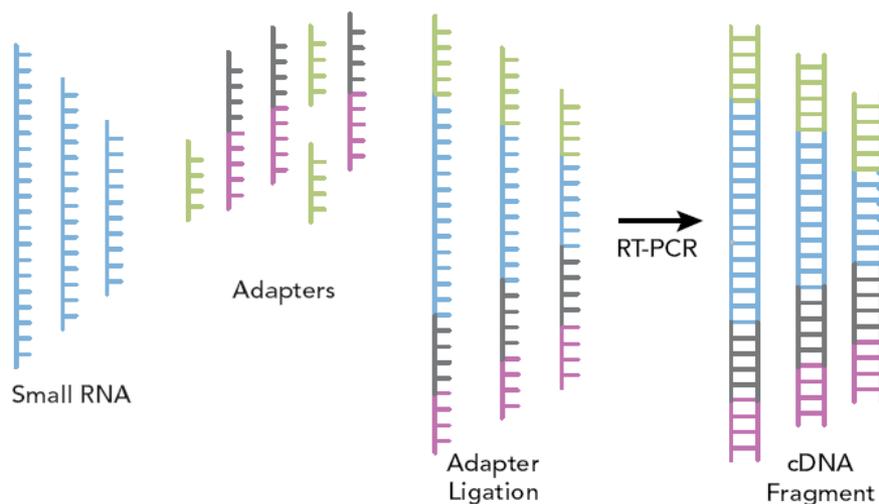
Here we used Illumina/Solexa sequencing for the identification of small RNAs in Chinese Hamster Ovary cells. This project was initiated and funded by the Consortium for CHO Cell Genomics and analysis was conducted at the University of Minnesota in cooperation with Prof. Wei-Shou Hu and the mammalian cell culture group.



**Figure 13 – Illumina/Solexa Sequencing Technology:** After ligation of 3' and 5' adaptors cDNA molecules are applied to the surface of single-molecule arrays or flow cells (1 and 2). There clusters of approximately 1000 clonal cDNA molecules are created by bridge amplification, where attached cDNA molecules bend over to form a bridge with adapters attached to the flow cell surface. For sequencing by synthesis one adaptor is cleaved leaving only equally orientated molecules in each cluster (not shown).

### 3.6.2. Small RNA Library Preparation

Preparation of cDNA libraries from total RNA was conducted at the Bioprocessing Institute in Singapore (BTI). Preparing the cell-free small RNA library started with isolating the small RNA fraction (18 -30nt) from total RNA followed by successive 3' and 5' adapter ligation, reverse-transcription to cDNA and PCR amplification (Fig. 14 shows a flow scheme).



**Figure 14 – small RNA library preparation for Illumina/Solexa Sequencing:** The library preparation consists of two ligating steps (3' and 5' adapters are ligated) and subsequent reverse transcription to cDNA. Finally (not shown) intact cDNA that have both adapters ligated are enriched using PCR reaction.

Briefly, isolated RNA aliquots were size fractionated using 15% Tris-borate-EDTA-urea-polyacrylamide (TBE-urea-gel) gels in order to isolate the small RNA fraction of 18 to 32 nucleotides in length. This isolation was followed by subsequent 5-prime sequencing adapter ligation, another TBE-urea-gel run, 3-prime adapter ligation and finally purification of 70 – 90 nucleotide sequences using a 10% TBE-urea-gel. In the next step RNA was reverse transcribed to cDNA using Super Script II Reverse Transcriptase at 44°C for 60 minutes and amplified in 15 cycles using PCR reaction. The PCR product was stored at 4°C. For extraction of fully amplified small RNA 50µl of PCR product were loaded onto a 6% TBE-urea gel. The band corresponding to 92bp was excised using a clean scalpel, and the gel fragment was dissolved by means of centrifugation through punctured sterile 0.5ml tube into a 2ml sterile tube. 100uL of 100x sterile 1X gel elution buffer were added to the gel debris and DNA was eluted for 2 hours under continuous rotation of the tube. The eluated DNA gel-debris mixture was centrifuged at full

speed for 2 minutes and 1uL of glycogen, 10uL of 3M NaOAc and 325uL of -20°C 100% ethanol were added followed by immediate centrifugation for 20 minutes at 14000 rpm. The supernatant was discarded and the remaining RNA pellet was washed using 500uL 70% ethanol. The dried pellet was finally resuspended in 10uL of resuspension buffer provided by Illumina. Quality parameters of all prepared libraries were assessed using Agilent RNA 6000 Nano Kit.

### 3.6.3. Small RNA Libraries

ID	Description (source of RNA)	Cell Line Characteristic
CHOK1	Parental CHOK1 - Adherent, serum-dependent - Harvested from Exponential Growth Phase	Parental Cell Lines
DG44	Parental DG44 - Adherent, serum-dependent - Harvested from Exponential Growth Phase	
DXB11	Parental DXB11 - Adherent, serum-dependent - Harvested from Exponential Growth Phase	
DG44MaB	Recombinant CHO DG44 (monoclonal Antibody) - Suspension adapted, serum-independent - Harvested from Exponential Growth Phase	Production Clone
CHOA_exp	Recombinant CHO DG44 (IgG) - Suspension adapted, serum-independent - Harvested from Exponential Growth Phase	CHO Aven temperature shifte / NaBu cultivation
CHOA_T33bu	Recombinant CHO DG44 (IgG) - Suspension adapted, serum-independent - Treated with 2 mM sodium butyrate during exponential growth phase 36 hours after low temperature (33°C) induction	
CHOA_T37	Recombinant CHO DG44 (IgG) - Suspension adapted, serum-independent - Harvested from Stationary Growth Phase	

Figure 15 – Summary of small RNA libraries prepared for Solexa Sequencing at BTI.

In total 7 small RNA libraries were prepared for Solexa sequencing and are briefly described in Figure 15. Three libraries were derived from parental cell lines (CHOK1, DG44 and DXB11, all serum dependent), one from a production cell line (DG44mAb, serum-free), and the remaining three libraries were prepared from a culture of CHO-Aven cells. Samples were taken from early exponential phase (CHOA\_exp), from temperature shifted (33°C) and sodium butyrate treated cells (CHOA\_T33bu), and late stationary phase (CHOA\_T37) at 37°C. Figure 15 gives a summary of all 7 small RNA libraries. Sequencing of all 7 libraries (each applied to a separate lane of the flow cell) was conducted at the National Center for Genomic Research (NCGR) in Santa Fe, NM.

#### **3.6.4. Data Processing**

Sequencing of the small RNA libraries yielded a set of 36-nt long sequences plus a corresponding quality score for each read. Our collaborators at NCGR filtered the data for low quality reads (such as repeat or poly(A) elements), and removed previously ligated sequencing adapters from each read, resulting in read lengths between 18 and 36 nucleotides (nt) (*high quality and trimmed reads*). During all subsequent alignment processes we maintained two FASTA files for each library, the first containing sequence information of only unique reads, while the second contained the corresponding information on read abundance.

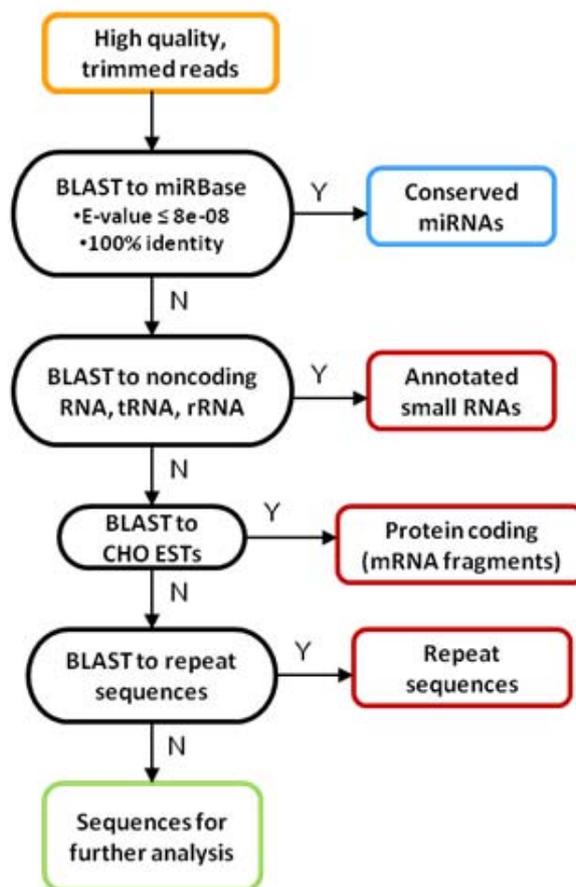
#### **3.6.5. BLAST Identification of conserved miRNAs**

For identification and annotation of conserved small RNA sequences from the Solexa dataset, TimeLogic Tera-BLASTX was used (Active Motif Inc, USA). BLAST alignments were set up on Linux computers of MSI (Minnesota Super Computing Institute) via a secure shell (SSH) connection using TimeLogic command syntax. For subtractive BLAST alignment the flow scheme depicted in Figure 16 below was developed. Tera-BLASTX alignment of Solexa reads to small RNA repositories was optimized by adjusting the BLAST parameter *extension size* to 20. Consequently, this algorithm was used to sequentially align Solexa reads to several small RNA databases. Table 9 below gives a summary of used databases and their online availability.

**Table 9 – Overview of small RNA databases**

Small RNA type	Description	Database
Piwi-interacting RNA (piRNA)	Regulation of germline development	piRNABank ( <a href="http://pirnabank.ibab.ac.in/">http://pirnabank.ibab.ac.in/</a> )
Transfer RNA (tRNA)	Amino acid delivery for protein translation	Genomic tRNA database ( <a href="http://lowelab.ucsc.edu/GtRNAdb/">http://lowelab.ucsc.edu/GtRNAdb/</a> )
Ribosomal RNA (rRNA)	Ribosome component; protein translation	SILVA rRNA databases ( <a href="http://www.arb-silva.de/">http://www.arb-silva.de/</a> )
Small nuclear RNA (snRNA)	Pre-mRNA splicing	NONCODE v.2* ( <a href="http://www.noncode.org">www.noncode.org</a> )
Small nucleolar RNA (snoRNA)	Modification of functional RNAs	NONCODE v.2
Small Cajal body associated RNA (scaRNA)	Modification of functional RNAs	NONCODE v.2
Repeat sequences	Non-coding repeat elements	RepeatMasker ( <a href="http://www.repeatmasker.org">www.repeatmasker.org</a> )

\*NONCODE also contains other noncoding RNAs and mRNA-like fragments



**Figure 16 – Subtractive BLAST alignment:** In the first step conserved miRNAs were identified by aligning reads to Sanger miRBase miRNA repositories, i.e. mature miRNA, miRNA\* and hairpin databases. After all hits had been removed from the Solexa dataset other non-coding small RNAs were identified using a variety of small RNA databases. Finally mRNA fragments and repeat sequences were annotated by BLASTing against a CHO EST database and repeatmasker database, respectively.

BLAST results in each step were filtered according to certain criteria (summarized in Table 10 below), in order to minimize the amount of false positive identifications. Stringency of BLAST criteria was adapted to target and query sequence length. E.g. for identification of mature conserved miRNAs higher significance values and percent identities were required as for annotation of repeats, or mRNA fragments. Only hits ranked as 1, referring to the top hit for each query sequence, were taken into consideration. For small RNAs known to have lengths between 18 and 32nt, i.e. mature miRNAs, miR\* and piRNA sequences we required the query sequence to be as long as or shorter than the target sequence (length constraint), which for example avoided hairpins being identified as mature microRNAs.

**Table 10 – Summary of BLAST Criteria used for small RNA identification**

BLAST Criteria	mature miRNA	miRNA*	Hairpin	piRNA	other small RNA databases
Rank	1	1	1	1	1
E-value	$8 \cdot 10^{-8}$	$8 \cdot 10^{-8}$	$10^{-6}$	$10^{-7}$	$10^{-6}$
Length Constraint	Yes	Yes	No	Yes	No
Percent Identity	100%	100%	? 90%	? 95%	<i>no restriction</i>

These criteria were set in Excel, and filtered results were exported to separate Excel result files. All hits were then removed from the initial Solexa dataset using in-house developed Perl scripts. Perl is a UNIX based script language combining features of C and shell scripting. It is primarily used for manipulating large textfiles or FASTA files and for developing simple algorithms for sequence analysis and manipulation.

After subtracting the all BLAST hits from the dataset, the reduced set was then used for the next BLAST alignment, following the flow scheme in Figure 16.

## 4. Results

### 4.1. MicroRNA Sequencing

#### 4.1.1. The Solexa Dataset

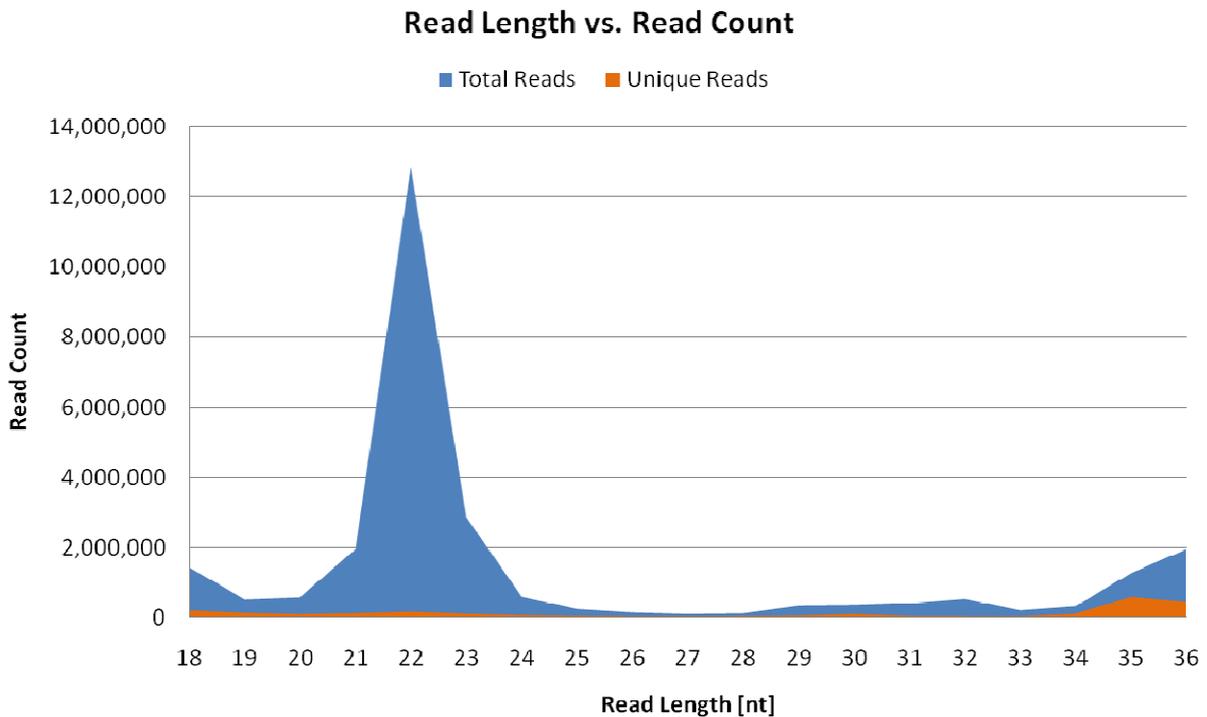
Since virtually no sequence information on CHO microRNAs or other small RNAs is publicly available so far, next-generation sequencing was applied on 7 CHO small RNA libraries as previously described (3.6.). Illumina/Solexa sequencing is based on parallel-sequencing of up to 40 million clonal cDNA clusters annealed to adapters on the surface of a flow cell. Thus, each sequenced cluster yields one small RNA sequence, commonly referred to as “read”. The total set of reads created by deep sequencing includes the sequencing results from all clusters on the Solexa/Illumina flow cell. Since a library or even a single cell can contain several copies of an RNA molecule, especially when it is biologically meaningful, it is possible that multiple clusters contain the exact same sequence; this sequence is then present several times in the *total Solexa data set*. In a first processing step our collaborators at NCGR removed these redundant sequences and created a *unique dataset* for each library, while the abundance information for each read was stored in a second file, referred to as the *abundance file*. As part of its quality control NCGR had removed all reads with low quality scores (e.g. poly(A) reads) and had trimmed sequence adapters resulting in a 11.7% loss of total reads. Thus, about 27 million sequences were left for BLAST alignment to various small RNA databases.

We received both files, the unique sequence file and the corresponding abundance file from NCGR and by multiplying each read with its abundance calculated how many reads had been initially present in the *total Solexa dataset*. Table 11 below summarizes the Solexa output for unique reads as well as total reads. Considering the ratio of total to unique Solexa reads (see Table 11) of 27 million to 2.7 million reads, Solexa sequencing yielded a 10x sequence coverage.

**Table 11 – Summary of Illumina/Solexa Data Output**

Library	Reads in the total data set	Reads in the total data set after filtering	Reads in the unique data set after filtering	Percent
CHOK1	3,267,033	2,961,582	275,966	90.70%
DG44	4,235,921	3,659,539	333,831	86.40%
DG44MaB	5,243,920	4,311,797	378,354	82.20%
DXB11	4,651,922	4,259,309	447,513	91.60%
CHOA_T37	4,729,775	4,256,586	416,410	90.00%
CHOA_T33bu	3,440,010	2,882,958	332,824	83.80%
CHOA_exp	5,073,483	4,736,997	557,072	93.40%
	30,642,064	27,068,768	2,741,970	88.33%

Due to the preparation technique used for the generation of cDNA libraries starting with excision of 18 – 30 nt bands (see section 3.6.2.) read lengths of the Solexa dataset varied between 18 and 36 nucleotides (up to 36 nt due to inexact band excision). We counted the frequency of each read-length to get an idea of how many sequences of a specific length were present in the Solexa dataset. Therefore the abundance/size distribution of unique reads as well as of total reads was analyzed and both distributions were superimposed in one graph (see Figure 17). Comparing the orange distribution for unique reads and the blue distribution for total reads, the great shift in frequency of 21 – 23 nucleotides long reads indicates the presence of numerous clonal copies for these read lengths. While the quality of the frequency values for these read lengths is assumed to be high, the results for lower (18 to 19bp) as well as higher (> 32bp) read length frequencies are less confident. Consequently the increase in frequency for 35 and 36bp reads is questionable. However, this data suggests that especially 22-23nt sequences are not random but biologically meaningful and that most probably mature miRNAs as well as other small RNAs are present in this CHO Solexa dataset.



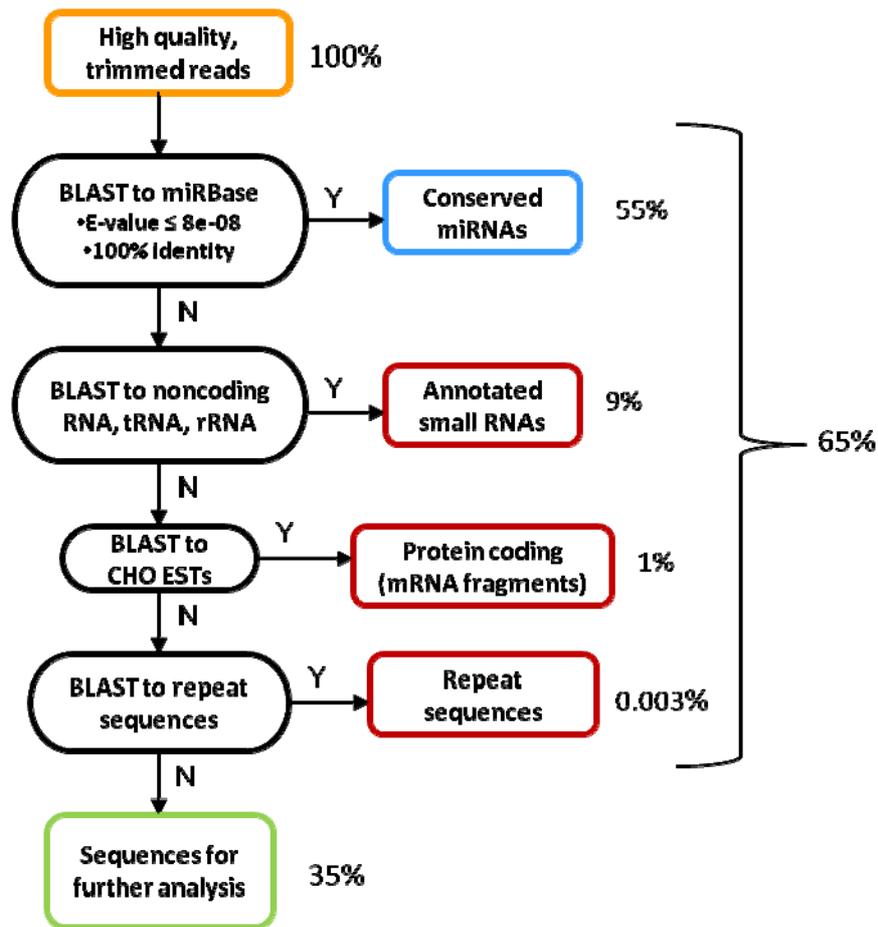
**Figure 17 – Comparison of read length versus read count for total (blue) and unique (orange) Solexa dataset:** Deep sequencing of small RNA in CHO picked up an enormous amount of 21 to 23 nucleotide long sequences. Yet, the amount of unique sequences of the same size is low, indicating that multiple copies of these sequences had existed in the sequencing libraries. Also for 18-nt as well as 31, 32, 35 and 36-nt sequences read count increases significantly from unique to total distribution.

#### 4.1.2. Annotation of Small RNAs

As described before, small RNAs were identified in this study based on their sequence conservation in humans or rodent species, such as mouse or rat, or in even further related species. In addition to *TimeLogic TeraBlast* (Active Motif Inc. Carlsbad, CA), other mapping algorithms were tested, such as SOAP (short oligonucleotide alignment program) or seqmap. Both mapping algorithms are commonly used to map short sequences such as Illumina/Solexa reads to longer reference sequences, e.g. messengerRNAs (Jiang and Wong 2008; Li, Li et al. 2008). However, BLAST alignment gave better results than SOAP or seqmap, which returned only few results, probably due to higher stringency. Thus, BLAST alignment in combination with subsequent filtering of BLAST results was chosen for the actual data analysis.

The focus of this study was to identify as many conserved microRNAs as possible; thus, BLAST alignment against mature miRNA, miRNA\* and hairpin databases was conducted first (using the complete Solexa dataset). According to the flow scheme in Figure 18, all microRNA hits were subtracted from the initial dataset. The remaining set of reads was subtractively aligned to several small RNA databases as well as mRNA and genomic-repeat databases, in order to narrow the amount of sequences for further analysis (e.g. novel CHO miRNA prediction) down.

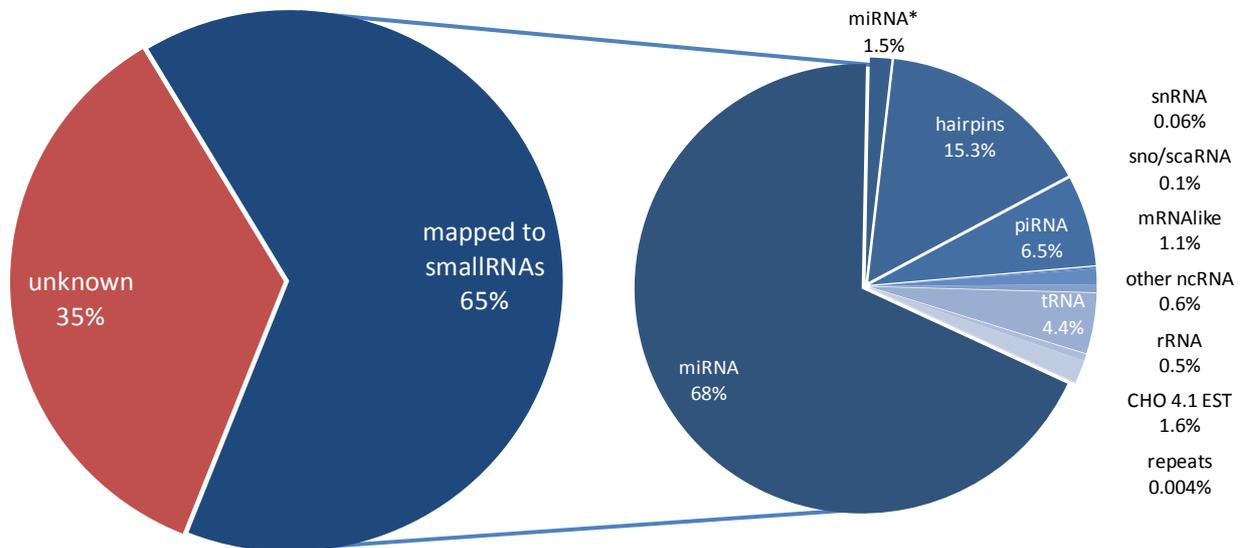
Mature miRNAs were identified as the most abundant class of small RNAs making up between 44% and 65% of total Solexa reads in 6 out of 7 libraries. Only the parental cell line DXB11 showed lower miRNA frequency, with only 9%. In general results obtained for DXB11 library did not follow the trend of the other 6 libraries, thus, DXB11 is assumed to be an outlier due to a bias induced during library preparation. On average about 1% of Solexa reads in each library were identified as miRNA\* sequences. This low frequency is explained by their fast degradation after the separation of miRNA:miRNA\* duplexes by the RISC machinery. BLAST alignment identified approximately 5% of reads as conserved miRNA hairpins. Only for DXB11 about 50% of Solexa reads were identified as hairpin miRNAs, again contradicting the general trend; presumably the effect of inexact band excision during library preparation. Among other non-coding small RNAs approximately 6.5% were identified as piwi interacting RNAs (piRNAs), which are a class of small RNAs of about 24 to 30 nucleotides in length. PiRNAs associate with Piwi proteins, which belong to the Argonaute superfamily and were found to be required for germline development in several species (Das, Bagijn et al. 2008; Wilczynska, Minshall et al. 2009). Transfer RNA (tRNA), ribosomal RNA (rRNA), small nuclear and nucleolar (sn/sno RNA) as well as small Cajal Body specific RNAs (scaRNA) were also identified and their distribution is given in Figure 19 below. Presumably, the reason for the low frequency of functional RNAs such as rRNA or tRNA is the short starting length for library preparation of approximately 18 bp to 32 bp. Thus, only fragments of tRNAs, 5S rRNAs and 5.8S rRNAs which have an average size of 80 nt, 120 nt and 160 nt, respectively, could be present in this dataset.



**Figure 18 – small RNA annotation strategy:** subtractive BLAST alignment for annotating conserved small RNAs in CHO. Next to each box of annotated small RNA class, the percentage of annotated reads (as part of the total initial number of reads) is given. In total 65% could be annotated, with microRNAs (including mature, star and hairpin sequences) being the most abundant class of conserved small RNAs. Other non-coding functional small RNAs include piRNA, tRNA and rRNA and make up 9% of the input data. Only few mRNA fragments (1%) and repeat sequences could be identified. For prediction of novel CHO miRNAs 35% of the initial dataset remain.

The pie chart in Figure 19 shows a detailed division of conserved small RNA reads, of which all microRNAs related sequences make up about 85%.

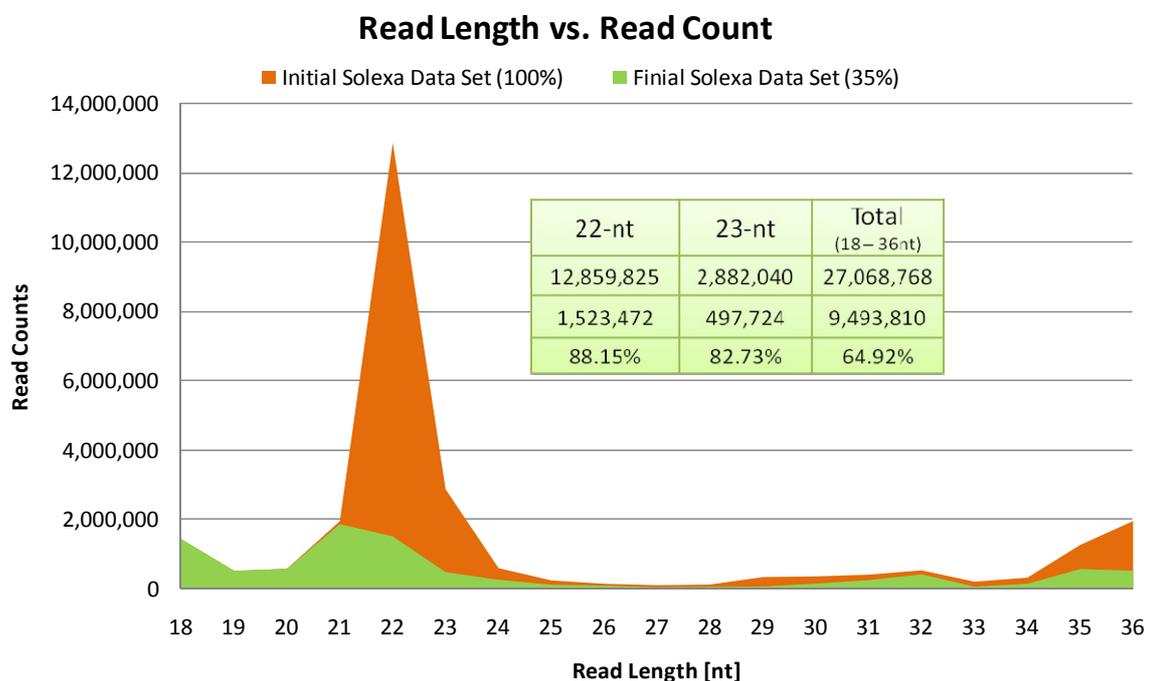
### Small RNA expression in CHO



**Figure 19 – Small RNA expression in CHO:** Among 65% of small RNA sequences that have been annotated in CHO based on sequence conservation, microRNA sequences are the most abundant ones: mature (68%), star (1.5%) and hairpin (15.3%) sequences make up 85%, followed by piRNAs (6.5%) and tRNAs (4.4%). Ribosomal RNAs (0.5%) and other non-coding RNAs (ncRNAs) make up only 1.1% of annotated reads. Also small nuclear and nucleolar RNAs (sn/snoRNA) were identified as well as small Cajal Body localized RNAs (scaRNA). CHO 4.1 EST and mRNA like sequences were identified based on alignment to a CHO EST database (version 4.1.) and noncode database, and refer to mRNA fragments that were picked up during library preparation.

After subtracting all BLAST results from the Solexa dataset about 35% of the initial 27 million reads (9.5 million) remained. This dataset was then analyzed for its size distribution and compared to the size distribution of the complete data set. Figure 20 shows the superimposition of both area plots: as can be seen, BLAST annotation of small RNA reads greatly reduced the 22-nt and 23-nt long sequences, leaving about 2 million of them unannotated, which corresponds to a reduction of more than 80%, and almost 90% for 22-nt sequences. Interestingly none of the shorter sequences (18-nt to 20-nt) could be mapped, whereas about half of the 35-nt and 36-nt sequences could be annotated. The inability to map shorter sequences might be due to low significance score of short alignments. Consequently, since all BLAST hits were filtered based upon these scores, short sequence hits might have escaped BLAST identification.

The great reduction of 22-nt sequences demonstrates that reasonable sequence conservation between hamster and other species, primarily human, mouse and rat, must exist. This high degree of conservation leaves only little space (10%) for CHO specific small RNAs of 22 nucleotides in length. However, this data only represents miRNA characteristics of Chinese Hamster Ovary Cells and CHO derived transformed cell lines. Consequently, not even one hamster tissue is represented by this data, and degree of conservation of the total set of Chinese Hamster miRNAs might be different.



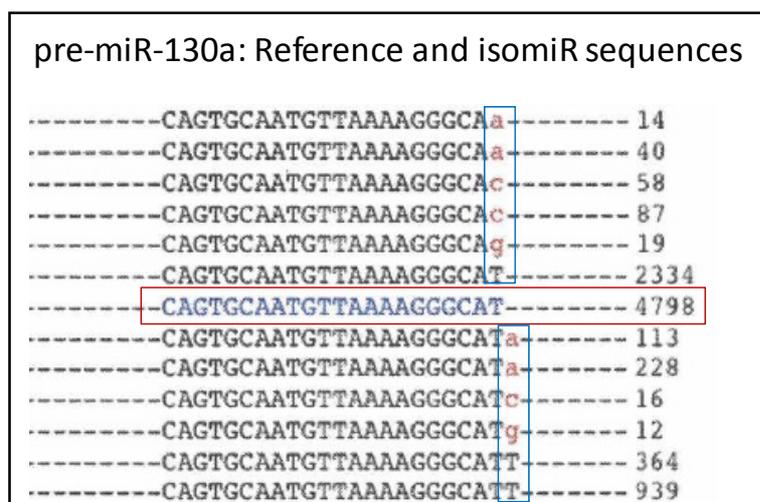
**Figure 20 – Comparison of size distribution before and after mapping small RNA sequences:** The orange curve corresponds to the previously shown read length/count distribution, representing the total Solexa dataset. The green curve represents the reduced dataset after several rounds of subtractive BLAST alignments. Overall 65% or 17.5 million reads were annotated leaving 9.5 million for further analysis. While the overall reduction amounted to 65%, more than 80% of 22-nt and 23-nt long sequences had been annotated.

#### 4.1.3. Annotation of Conserved MicroRNAs

##### 4.1.3.1. Variability in microRNA processing - isomiRs

In total 337 mature CHO miRNAs were identified, and, for almost every mature miRNA sequence several isomiR sequences were detected. When all isomiR sequences are considered,

an overall of 15,367 sequences had been annotated as mature miRNAs, resulting in an average of 45 isomiRs per microRNA. IsomiR sequences have first been detected during deep sequencing of small RNAs (Morin, O'Connor et al. 2008), and are characterized by slight sequence variation from their reference sequence, predominately at their 3' ends ("reference sequence" refers to the sequence that is stored in the Sanger miRBase repository). Figure 21 shows characteristic isomiR sequences for human microRNA 130a.



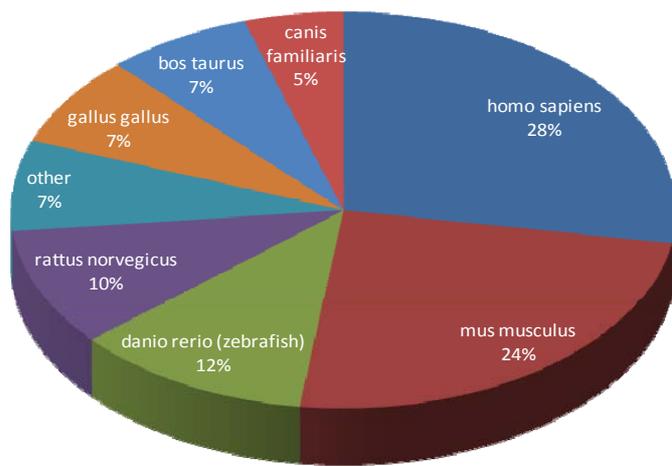
**Figure 21 – Characteristics of isomiR sequences:** Depicted is a snapshot of characteristic isomiR sequences from Morin, O'Connor et al. (2008). Most variations are the consequence of inexact Drosha and Dicer cleavage, leading to single nucleotide extensions or reductions. The transparent red box indicates the Sanger reference sequence for miR-130a, while blue boxes highlight single nucleotide extension and reduction.

It is supposed that isomiRs are the result of imprecise enzymatic processing - either Dicer1 or Drosha - or of single nucleotide extension during miRNA biogenesis, which is most commonly observed on the 3'-end but in some cases also at the 5'-end. Interestingly, some isomiR sequences exhibited high abundances that were far above insignificant background abundances; and for some miRNAs not the read identical to the miRBase reference sequence, but an isomiR sequence showed highest abundance. This makes the choice of miRNA for differential expression analysis difficult, since studies have been shown that most abundant isomiR and miRBase sequence can give different results for differential expression analysis (Morin, O'Connor et al. 2008). Due to legal issues no isomiR sequences on the example of CHO

sequences can be given. Thus, Figure 21 shows characteristics of isomiR sequences for miR-130a, taken from a study by (Morin, O'Connor et al. 2008).

#### 4.1.3.2. Characteristics of microRNA Conservation in CHO

15,367 isomiRs or 337 distinct conserved microRNAs in CHO cells were identified based on sequence alignment with miRNAs coming from altogether 25 species.



**Figure 22 – Conserved mature miRNAs were identified in several species:** 25 species were shown to contain miRNA sequences homologous to CHO, with >60% coming from human, mouse and rat. Also zebrafish, chicken, cow and dog yielded considerable results. Among the 18 species referred to as “other” gorilla, platypus, frog, worm and fruit fly were present.

Figure 22 depicts the distribution of species in a pie chart, which shows that the majority of the 337 sequences (>50%) was identified based on conservation in Homo sapiens and Mus musculus (house mouse). Presumably, this is due to the greater availability of mature miRNA sequences for these species. In some instances it had occurred that reads with high sequence similarity had hit the same miRNA in different species. For example, in case BLAST alignment had identified a read as “hsa-miR-21” and another similar read (e.g. isomiR) as “bta-miR-21”, both hits were reduced to “miR-21” and subsequently counted as one hit. This way the originally 337 miRNA sequences were narrowed down to a total of 260 unique conserved miRNAs.

Table 12 summarizes the results obtained from conserved miRNA annotation. While in total 260 conserved miRNAs were identified, not every library contained all of these. Therefore the third column in Table 12 (number of unique miRNAs) lists how many unique miRNAs out of 260 were contained in every library. Across all 7 libraries a common set of 147 miRNAs was identified. In case a miRNA was not picked up in more than 50% of libraries (4 or more) we denoted it as “singleton” expressing less confidence.

**Table 12 – Summary of conserved miRNA identification**

Library	Number of reads after data filtering	Number of miRNA reads	Number of unique miRNAs	Range of miRNA abundance	Median of abundance	Most abundant miRNA	Number of Singletons <sup>1</sup>
CHOK1	2,961,582	1,505,248	196	1 to 793,110	70	let7f	9
DG44	3,659,539	1,610,826	198	1 to 702,428	110	let7f	8
DG44MaB	4,311,797	2,785,876	187	1 to 1,469,320	67.5	let7f	3
DXB11	4,259,309	400,361	179	1 to 84,525	55	let7c	7
CHOA_T37	4,256,586	2,125,922	216	1 to 949,805	79	let7f	9
CHOA_T33bu	2,882,958	1,433,791	212	1 to 689,508	38.5	let7f	11
CHOA_exp	4,736,997	2,167,187	221	1 to 905,628	73	let7f	12

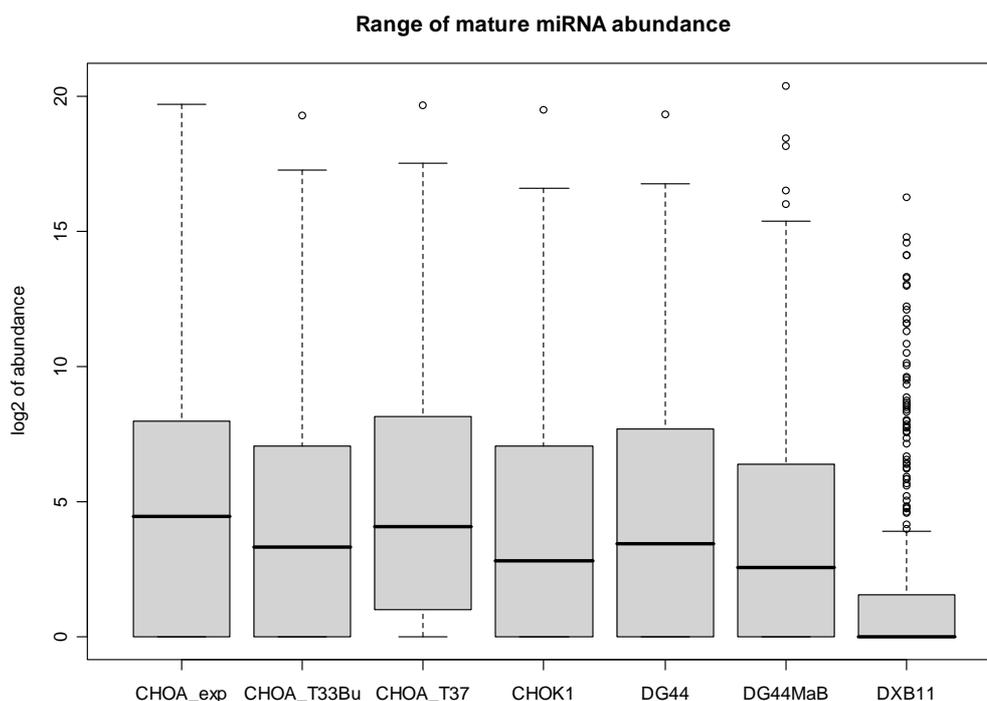
<sup>1</sup> Singleton: miRNAs that aligned to 3 or fewer sequences across all 7 libraries

The great range of miRNA abundance - from mRNAs with 1 single copy up to > 1 million copies per library for miR-let7f - indicates that microRNA expression undergoes great dynamics in CHO cells.

#### 4.1.4. Mature microRNA Expression Analysis

Abundance distribution for all 337 mature microRNAs in CHO was assessed using descriptive boxplot statistics in R, in order to compare unnormalized mature miRNA frequencies throughout the 7 libraries (Figure 23). Therefore the log<sub>2</sub> of abundance was calculated to allow better scaling. Except DXB11 all libraries show relatively uniform results, with the median of frequency ranging between 6 and 22 read counts. Also, this analysis identifies DXB11 as an outlier and indicates that CHO Aven derived libraries are characterized by slightly higher median frequencies than the parental cell lines or DG44mAb.

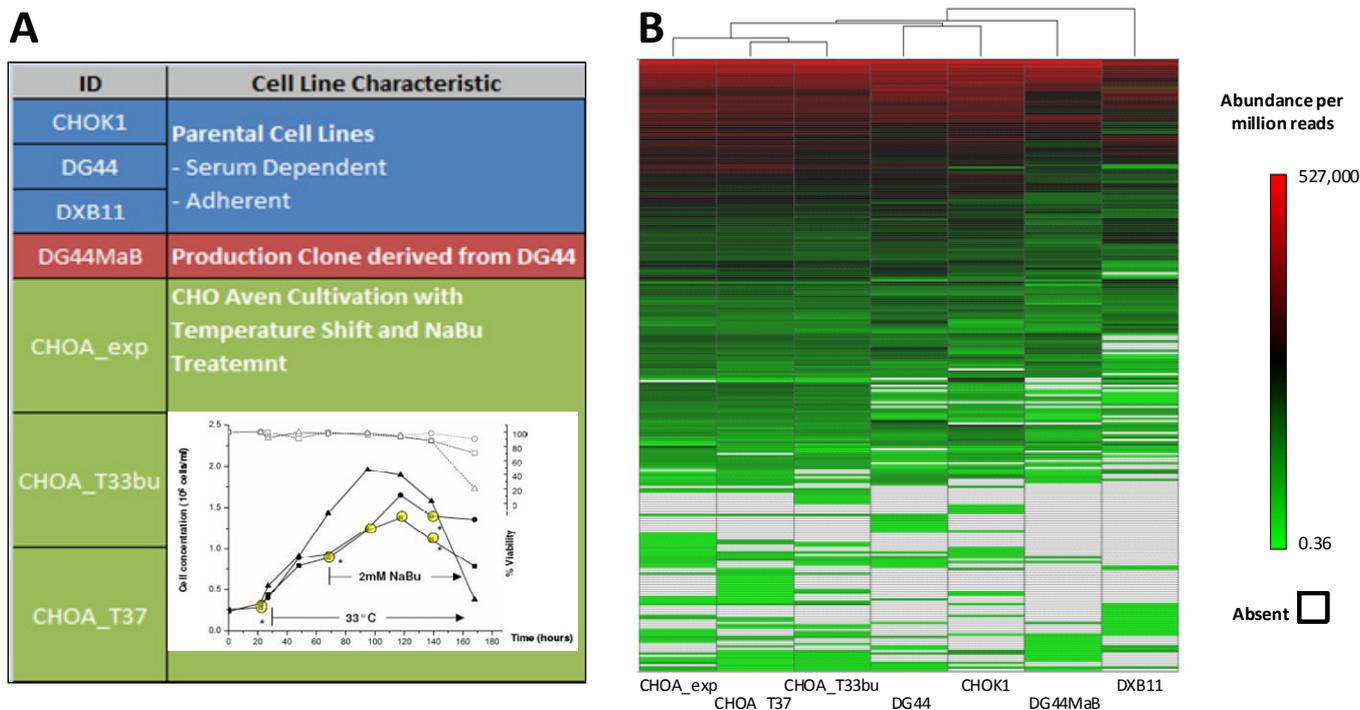
Boxplot Statistic	CHOA_exp	CHOA_T33Bu	CHOA_T37	CHOK1	DG44	DG44MaB	DXB11
Extreme of Lower Whisker	1	1	1	1	1	1	1
Lower Hinge	1	1	2	1	1	1	1
Median	22	10	17	7	11	6	1
Upper Hinge	256	136	287	134	206	84	3
Extreme of Upper Whisker	872,210	158,829	191,021	100,323	111,974	43,620	15
Number of Outliers	0	1	1	1	1	5	60
Value of highest outlier	n/a	636,862	847,188	744,055	662,263	1,387,968	79,315



**Figure 23 – Descriptive Boxplot statistics of mature miRNAs expression:** Box plots of mature miRNA frequency and Boxplot statistics above indicate differences between DXB11 and the other libraries, suggesting errors during DXB11 library preparation.

In order to use the abundance information to visualize similarities and differences in terms of miRNA expression between the seven libraries, a heat map was created and hierarchical clustering was performed. Since the initial number of reads was different for each library, the abundance data had to be normalized. Therefore each miRNA abundance value was divided by the number of million reads in that library, resulting in normalization as “abundance per million reads”. These values were then  $\log_{10}$  transformed and inserted in Spotfire Software, which was used for creating a heatmap and hierarchical clustering of libraries that is shown in Figure 24. Red color corresponds to high expression, green to low expression and grey color indicates absent microRNAs in the heatmap. The abscissa has all 7 libraries listed, while the ordinate

consists of all 260 unique miRNAs. Clustering of libraries according to their miRNA expression profile confirms that DXB11 is an outlier, and therefore DXB11 is not further considered in this analysis.



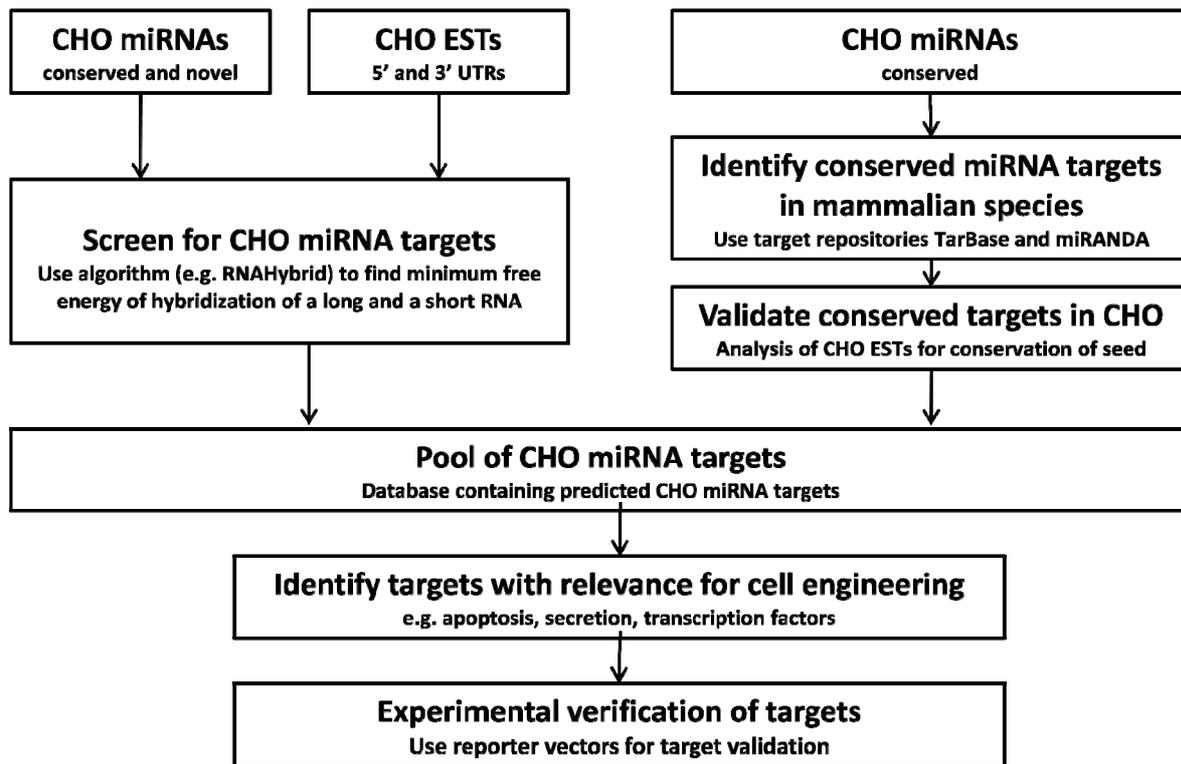
**Figure 24 - miRNA expression levels:** (A) Brief description of small RNA libraries used for RNA-Seq. Three parental cell lines, CHOK1, DG44 and DXB11; one production clone, DG44MaB; three samples derived from CHO Aven Batch Cultivation similar to the depicted growth curve. (B) Illumina/Solexa abundance values are shown normalized as transcripts per million reads and  $\log_{10}$  transformed. Red: high expression; black: intermediate expression; green: low expression; grey: absent in this library.

Hierarchical clustering of miRNA expression shows that CHO Aven derived samples (exponential phase/stationary phase/temperature shift + NaBu treatment) exhibit a high degree of similarity. Although on higher levels parental cell lines DG44 and CHOK1 cluster, and finally DG44MaB, it is obvious that differences between distinct cell lines are higher than within one cell line under different culture conditions (CHO Aven). This observation suggests that changes in global miRNA expression patterns are rather caused by profound genetic changes such as stable transfection of a gene cassette plus subsequent gene amplification, than by exposing cells to serum-free adaption, cold-shock conditions or treatments with small molecule enhancers (sodium butyrate). This is also confirmed by the fact that the DG44 derived production cell line,

DG44MaB clusters later with its parental cell lines CHOK1 and DG44 than CHO Aven cells. However, differential expression of microRNAs upon a change in external parameters has been shown before and yielded potential microRNA candidates for cell line engineering (Gammell, Barron et al. 2007).

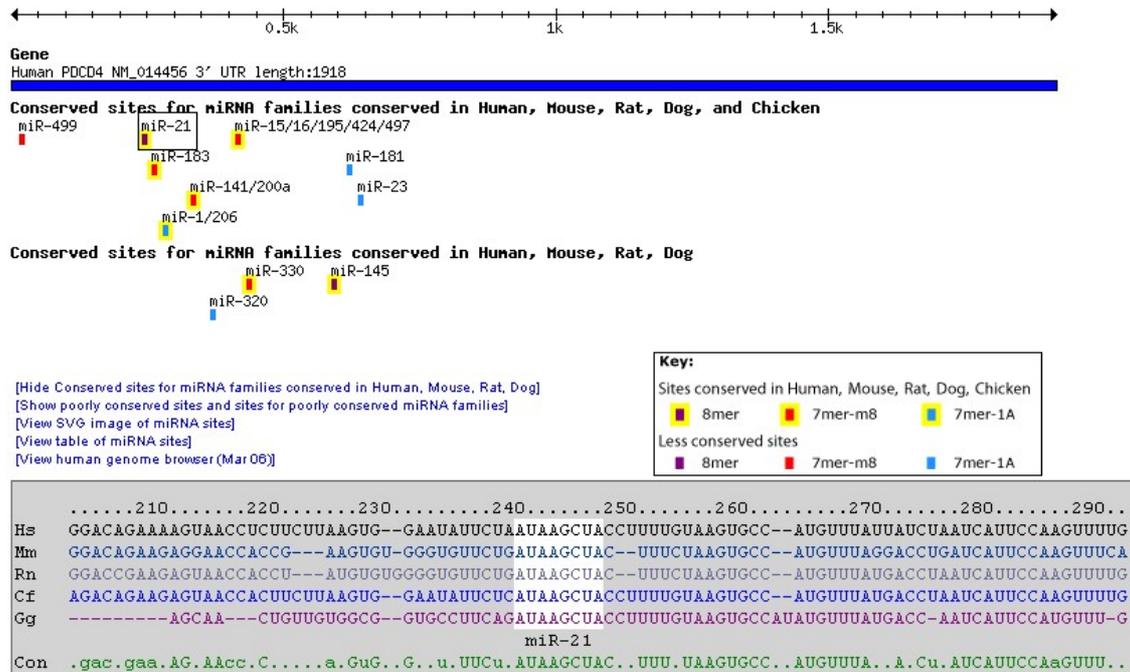
#### **4.2. MicroRNA Target Prediction in CHO Cells**

In order to be able to specifically use microRNAs for improving cell phenotypes it is crucial to know about their regulatory functions, i.e. which mRNAs are targeted and down-regulated by a specific microRNA. Therefore a strategy was developed to use the now available knowledge of CHO microRNA sequences for CHO miRNA target prediction. We came up with two approaches that can be pursued in parallel (see Figure 25): first all known CHO miRNA sequences, either identified based on conservation or bioinformatic prediction, can be combined with CHO transcriptome data (here especially 3' and 5' UTR sequences are of high interest) to run prediction algorithms such as RNAHybrid. This algorithm calculates the minimum free energy of hybridization between microRNAs and a potential target sequences and yields a set of putative target mRNAs. However, this approach for target identification requires in depth bioinformatic knowledge and experience and was not pursued in this study. A second strategy for CHO microRNA target identification relies upon cross-species conservation of miRNA targets: several public databases summarize current knowledge about microRNA targets and target conservation e.g. tarbase, targetscan or miRANDA (John, Enright et al. 2004; Lewis, Burge et al. 2005; Sethupathy, Corda et al. 2006). These databases provide miRNA seed regions as well complementary 3'UTR target sites in messenger RNAs.



**Figure 25 – Flow Scheme for miRNA target identification.** Two approaches can be pursued in parallel. However, due to lower demands on bioinformatic knowledge, only the strategy shown on the right side was pursued.

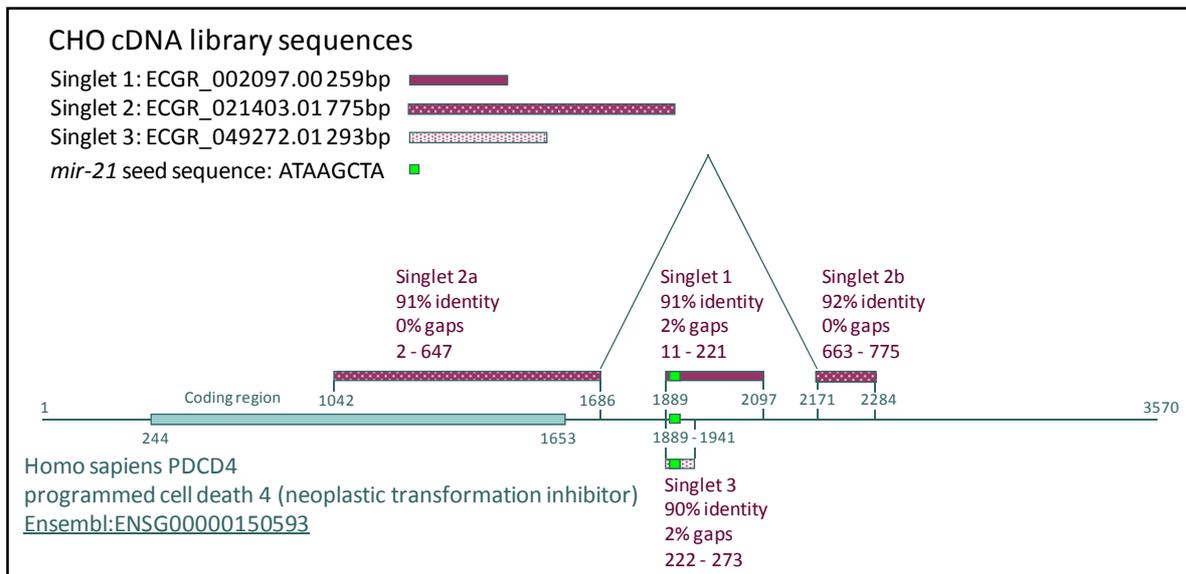
In case of PDCD4, Zhu et al. published the miR-21 seed region as well as the complementary target site in human PDCD4 3' UTR in 2008 (Zhu, Wu et al. 2008). This stretch in the PDCD4 3' untranslated region was identified as highly conserved across species using targetscan. Figure 26 shows a screenshot of the results from a TargetScan search for miRNAs binding human PDCD4 3'UTR. As can be seen binding sites for several miRNAs are available, but only one 8mer binding site which additionally exhibits strong cross-species conservation. The lower section of Figure 26 shows a comparison of human, mouse, rat, dog and chicken PDCD4 3'UTR sequences. According to this data the binding site within PDCD4 to the miR-21 seed region is highly conserved. Thus, it makes sense to search for the exact same sequence in CHO PDCD4.

**Human PDCD4 3' UTR**


**Figure 26 – Potential miRNA binding sites in human PDCD4 3'UTR:** a miR-21 binding site with perfect 8mer complementary to the miR-21 seed region is located around nucleotide 240. The sequence alignment in the lower section shows that the miR-21 binding site is highly conserved in human, mouse, rat, dog and chicken. Thus, as a further rodent species, also Chinese Hamster might bear this conserved binding site.

Therefore Chinese Hamster singleton sequences annotated to human and mouse PDCD4 were retrieved from CHO Consortium CHO EST database (access via ACBT – Sandoz/Novartis) and BLAST aligned to human PDCD4 (RefSeq. ID NM\_011050.1|XM\_983657.1) using BLAST2Seq. As can be seen in Figure 27 below, three CHO singlets could be aligned to human PDCD4. Of singlet 2, nucleotides 2 to 647 (denoted as *singlet 2a*) align to parts of the PDCD4 coding region and 3'UTR (nucleotides 1042 – 1686), whereas nucleotides 663 to 775 (denoted as *singlet 2b*) align to a more distant part of PDCD4. This suggests the existence of an alternative splice site within CHO PDCD4. Within this alternatively spliced region, another CHO singleton (CHO PDCD4 singlet 1) aligns to human PDCD4, nucleotides 1889 – 2097. This is the region on human PDCD4 that

contains the published miR-21 seed sequence, ATAAGCTA – which is also present on CHO PDCD4. The third CHO PDCD4 singlet (denoted as singlet 3) also contains the miR-21 seed sequence and aligns to this part of the human PDCD4 3'UTR region.



**Figure 27 – Blast 2 sequence Alignment of CHO PDCD4 ESTs to human PDCD4:** Conserved miR-21 binding site exists in CHO PDCD4 gene, but might be subject to alternative splicing.

Based on these results and due to the publicly available evidence that cgr-miR-21 is identical to hsa-miR-21 (Gammell, Barron et al. 2007) it was decided to study the effects of miR-21 overexpression and knockdown in a CHO production cell line.

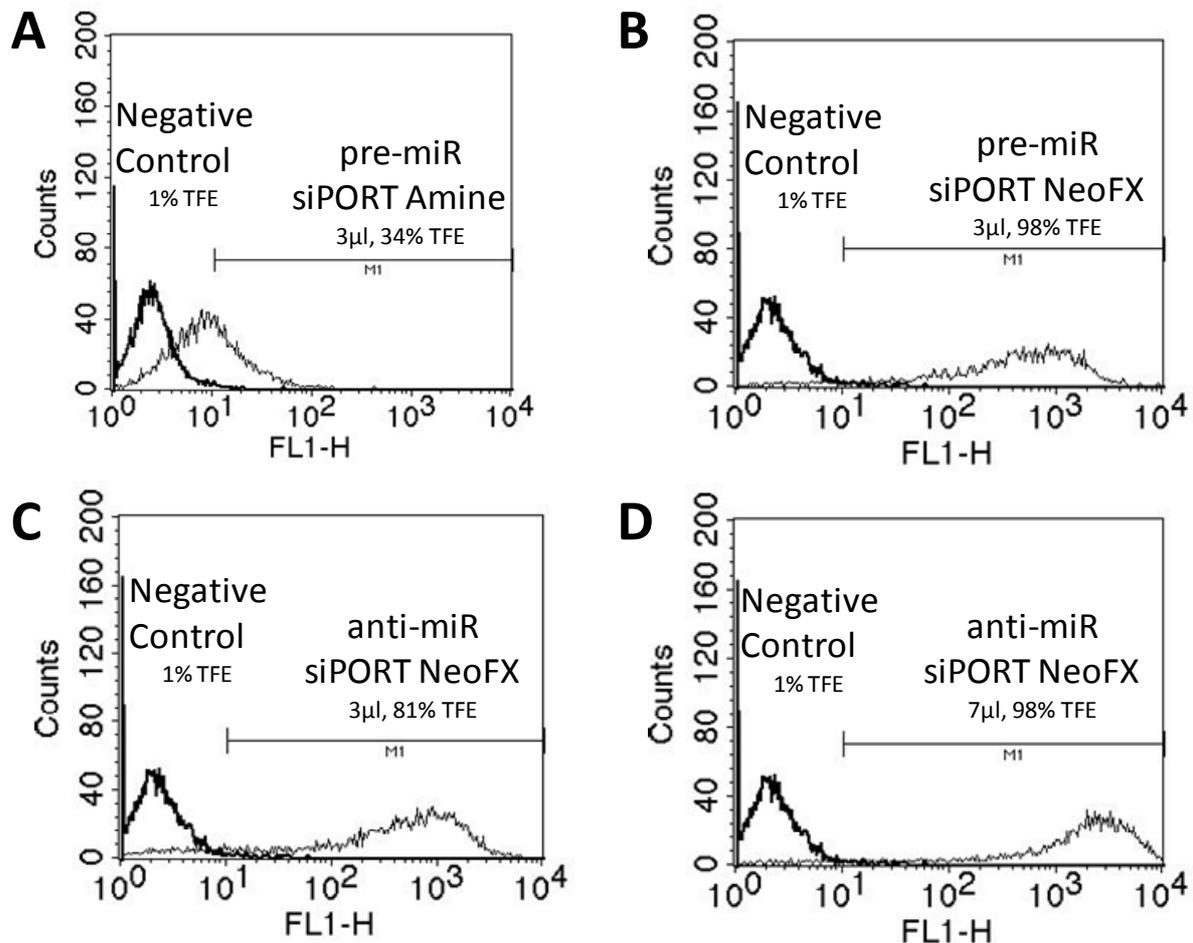
### 4.3. MicroRNA Overexpression and Inhibition in CHO

#### 4.3.1. Optimization of transient microRNA Transfection

The most important factor influencing transient transfection efficiency (TFE) is the transfection agent. According to the manufacturer, factors like RNA-concentration or cell density are of less importance as long as cells show high viability at the time of transfection. Ideally optimized transient transfections are characterized by high transfection efficiencies combined with low cytotoxic effect on the cells.

For this study two different transfection agents were tested, siPORT Amine and siPORT Neo<sup>FX</sup>. Optimization of kind and amount of transfection agent was planned based on a one factor

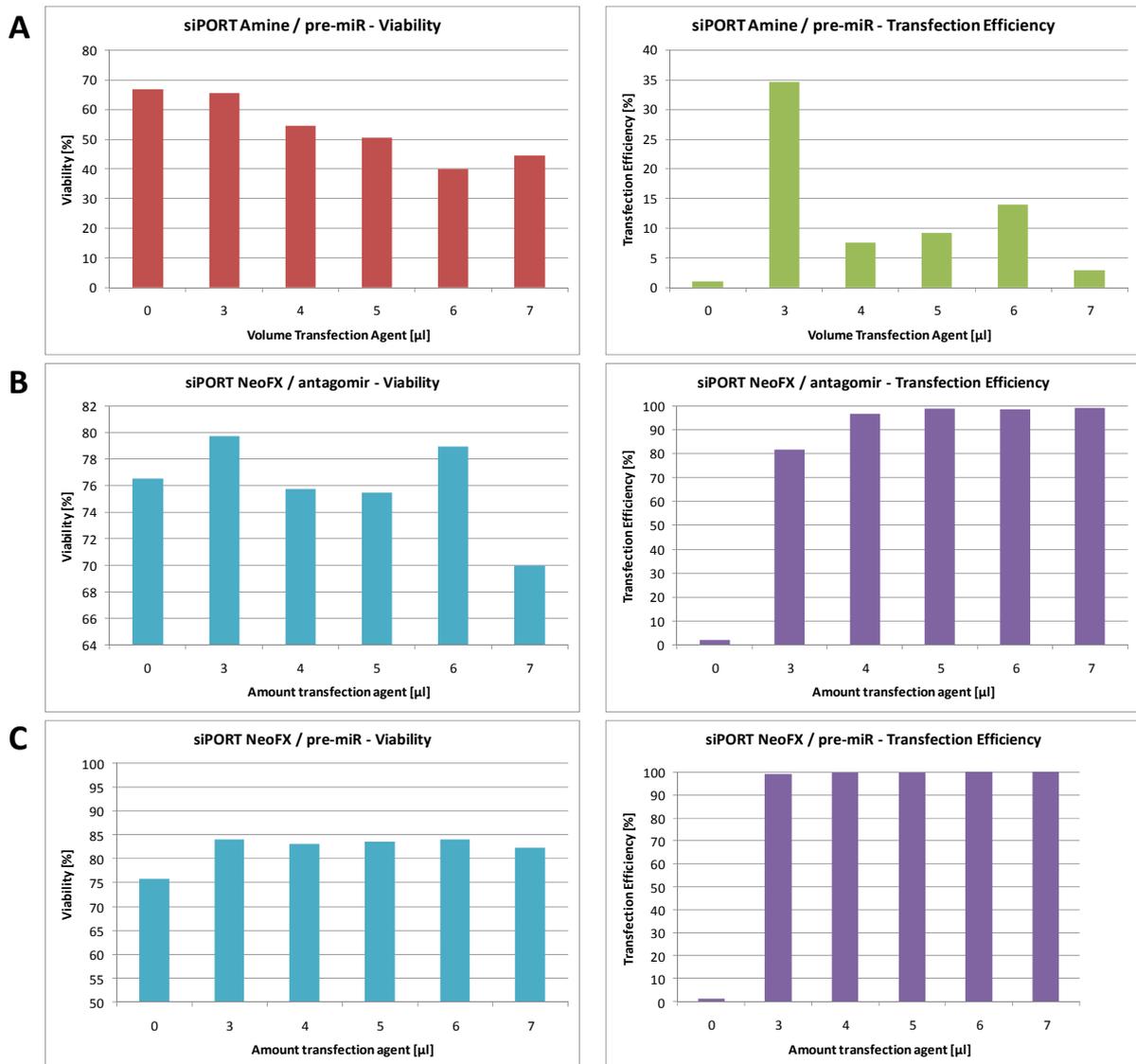
response surface method, using Design Expert (Stat-Ease Inc., Minneapolis, MN). Both transfection agents were tested in a range of 3 – 7  $\mu$ l, while the amount of RNA (30 nM) and cell concentration ( $1 \times 10^5$  cells/ml) were kept constant.



**Figure 28 – Flow Cytometry histograms of optimization of transfection efficiencies:** A: histogram indicating a transfection efficiency (TFE) of 34% for 3  $\mu$ L siPORT Amine transfection agent together with precursor miRNA molecules. B: in comparison to A, TFE can be improved to 98% using a different transfection agent – siPORT Neo<sup>FX</sup>. C: While 3  $\mu$ l siPORT Neo<sup>FX</sup> yield high TFEs for miRNA precursors (see B), for antagomir (anti-miR) transfection only a TFE of 81% is achieved. D: increase in the amount of siPORT Neo<sup>FX</sup> transfection agent to 7  $\mu$ l improves TFE from 81% to 98%.

As can be seen from both, flow cytometry raw data in Figure 28 as well as the summarized results in Figure 29, siPORT Amine transfections were characterized by poor performance, exhibiting low transfection efficiencies as well as low viabilities, going down to 40% for volumes of 6  $\mu$ l and 7  $\mu$ l. However, reasonable results were obtained for transfection of both, precursor

miRNAs as well as antagomirs, when siPORT Neo<sup>FX</sup> was used. Figure 28 shows a significant shift in FL-3 intensities from negative control samples to transfected samples, indicating high overall transfection efficiencies. Hence, TFEs of more than 95% were constantly observed for siPORT Neo<sup>FX</sup> transfections, with the exception of the 3  $\mu$ L transfection of antagomirs.



**Figure 29 – Comparison of viability and transfection efficiency for FAM-pre-miRNA.** A: siPORT Amine transfection results in terms of viability (red) and transfection efficiency (TFE, green). B,C: in comparison to siPORT Amine, siPORT Neo<sup>FX</sup> yields higher transfection efficiencies of both precursor miRNAs (B) and antagomirs (C) while maintaining cell viability at a higher level. The level of viability is generally low, which is not a result of siPORT Neo<sup>FX</sup> cytotoxicity, since also untransfected controls (0  $\mu$ L bars) exhibit

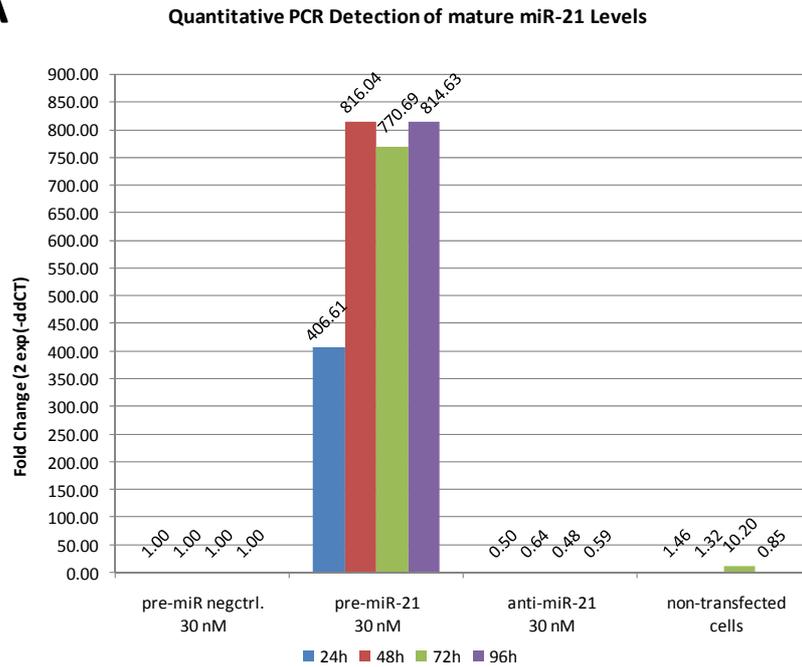
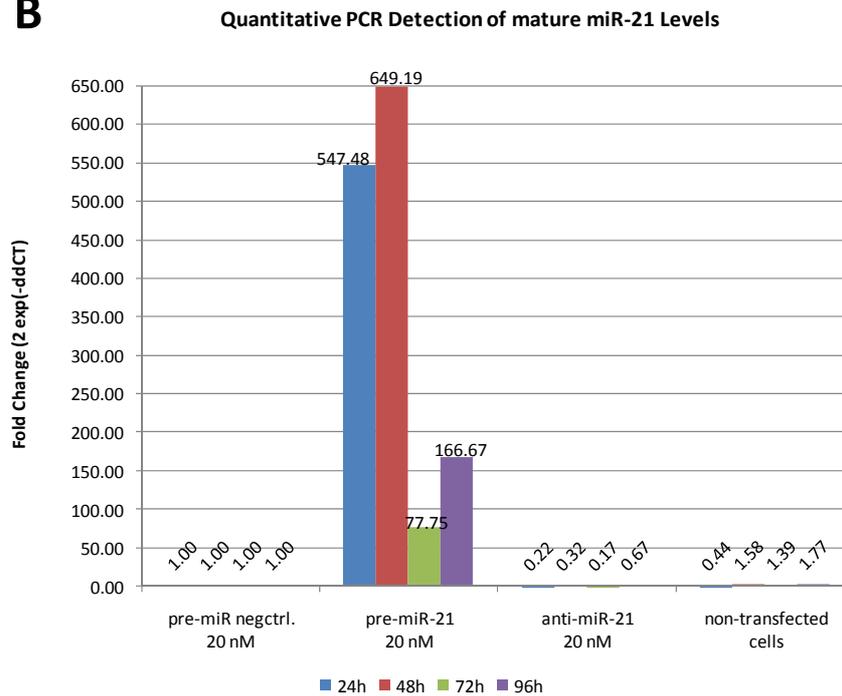
low viability. It is assumed that PBS washing steps of cell pellets prior to FACS analysis exerted stress on the cells and increased cell death.

Based on these results, 5  $\mu$ l siPORT Neo<sup>FX</sup> were used of future transient transfections of both, precursor miRNAs and antagomirs. Due to the already high transfection efficiencies as well as high FAM signals (not shown), RNA concentration (30 nM) and cell density ( $10^5$  cells/ml) were not further optimized. While these results confirm the intracellular uptake of fluorescently labeled Ambion miRNA precursor and antagomirs, intracellular uptake and subsequent induction of mature miR-21 levels as well as effective inhibition of mature miR-21 by antagomirs had to be confirmed.

#### **4.3.2. Real-Time qPCR Analysis of miR-21**

In order to confirm correct processing of pre-miRNAs to mature miRNAs as well as inhibition of endogenous miR-21 by anti-miR-21, Taqman real-time PCR studies of pre-miR-21 / anti-miR-21 transfected samples were performed. For analysis the levels of mature miR-21 were normalized to the amount of U6 snRNA in each sample (deltaCt) and in the next step, related to the negative control transfection for each time-point. Hence, fold changes for nc-pre-miR are 1 for every time-point.

As can be seen in the upper diagram of Figure 30, the use of 30 nM RNA concentrations resulted in a great induction of mature miR-21 in CHO cells, with a 400-fold induction after 24h and an 800-fold induction after 48h. This level of induction remained constant for at least another 48 hours. At the same time, transient transfection of miR-21 antagomirs resulted in a 50% knockdown in mature miR-21 levels throughout 96 hours of cultivation. For 20 nM RNA concentrations different miRNA inductions are observed (lower diagram in Figure 30): while the initial levels of miR-21 overexpression are almost equally good, resulting in a 650-fold induction after 48 hours, miR-21 levels decrease rapidly to a  $\sim$  160-fold induction after 96 hours. Inhibition of endogenous miRNAs with anti-miR-21 resulted in a constant down-regulation of 80% for 72h. After 96 hours miR-21 knockdown had decreased to 40%. Yet, the high deviation in values for untreated controls shows that the qPCR is prone to high standard deviations probably due to handling issues. Thus, all further experiments based on miR-21 transfection and inhibition required individual qPCR analysis.

**A****B**

**Figure 30 – Taqman qPCR:** Intracellular amounts of mature miR-21 during a 96h batch cultivation (batch run #1 and #2) were analyzed relative to the miR-21 levels of a non-functional precursor (nc-pre-miR) transfection. A,B: the bar chart in A depicts results for a 30 nM RNA transfection, while the bar chart in B represents 20 nM RNA concentrations. High inductions of miR-21 levels for miR-21 precursor samples as

well as constant knockdown of miR-21 for antagomir samples were observed, which confirm the validity of batch cultivation results from section 4.3.3.

#### 4.3.3. Batch Cultivation of miR-21 Transfected CHO Cells

Batch cultivations of transiently transfected cells were performed for 96 hours with constant monitoring of cell density, viability and EpoFc product titers. Two independent batch cultivations in 6-well plates were conducted, and are referred to as run #1 and run #2.

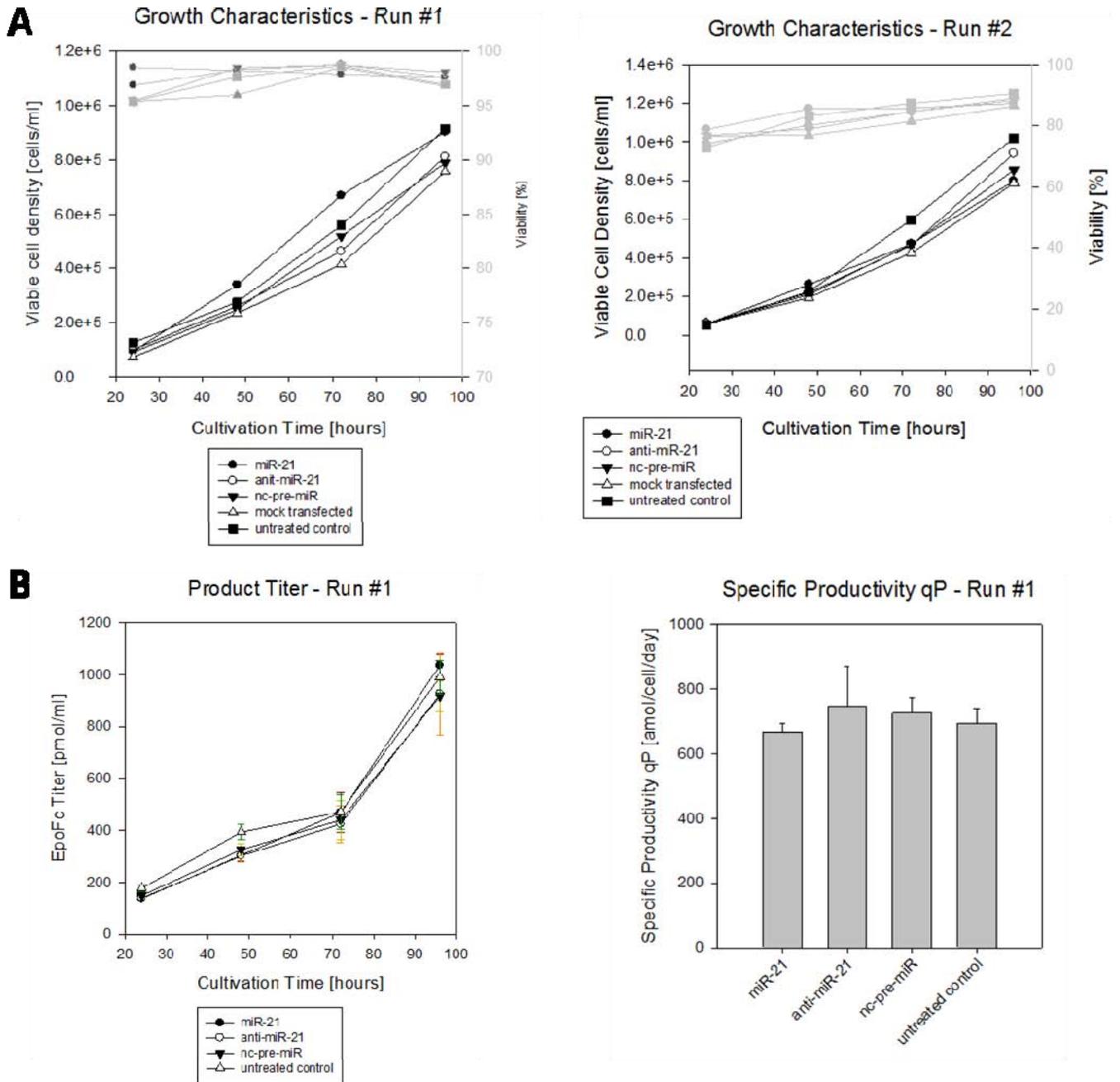
##### 4.3.3.1. Cell Growth

After an initial 24h lag phase following transfection, cells entered growth phase with growth rates ranging between 0.50 and 0.58 d<sup>-1</sup>. Table 13 below summarizes growth rates for each sample and batch cultivation.

**Table 13 – Growth rates for Batch cultivation #1:**

Sample	Run #1	Run #2
	Growth rate $\mu_{0-96h}$ [d-1]	Growth rate $\mu_{0-96h}$ [d-1]
miR-21	0.550	0.519
anti-miR-21	0.523	0.560
nc-pre-miR	0.516	0.537
mock transfected	0.506	0.515
untreated control	0.553	0.581

While untreated cells exhibited the highest growth rate in both runs, mock transfected cells were characterized by the slowest growth rate. Transfected cells did not show specific patterns of high or slow growth rates compared to negative control samples. Figure 31 below depicts growth characteristics as well as product titers for batch run #1. After 96 hours cell densities of  $1 \cdot 10^6$  cells/ml were observed for untreated cells. Although miR-21 transfected samples showed fastest growth and highest cell densities 72 hours after transfection, this effect was not stable and could neither be detected after 96 hours nor in batch run #2. Decrease in mature miR-21 levels cannot be responsible for this observation, since qPCR control indicated high miR-21 induction for at least 96 hours using 20 nM and 30 nM RNA concentrations for transfection.



**Figure 31 – Batch cultivation of miR-21 transfected cells:** (A) Viable cell density and viability as functions over cultivation time are shown for both runs. (B) For batch run#1, EpoFc product titers were measured throughout the cultivation. The left graph shows EpoFc product concentration in pmol/ml over time. After calculation of cumulative viable cell days (CD, not shown) specific productivities for each transfection was calculated and is illustrated as a bar chart in the right graph.

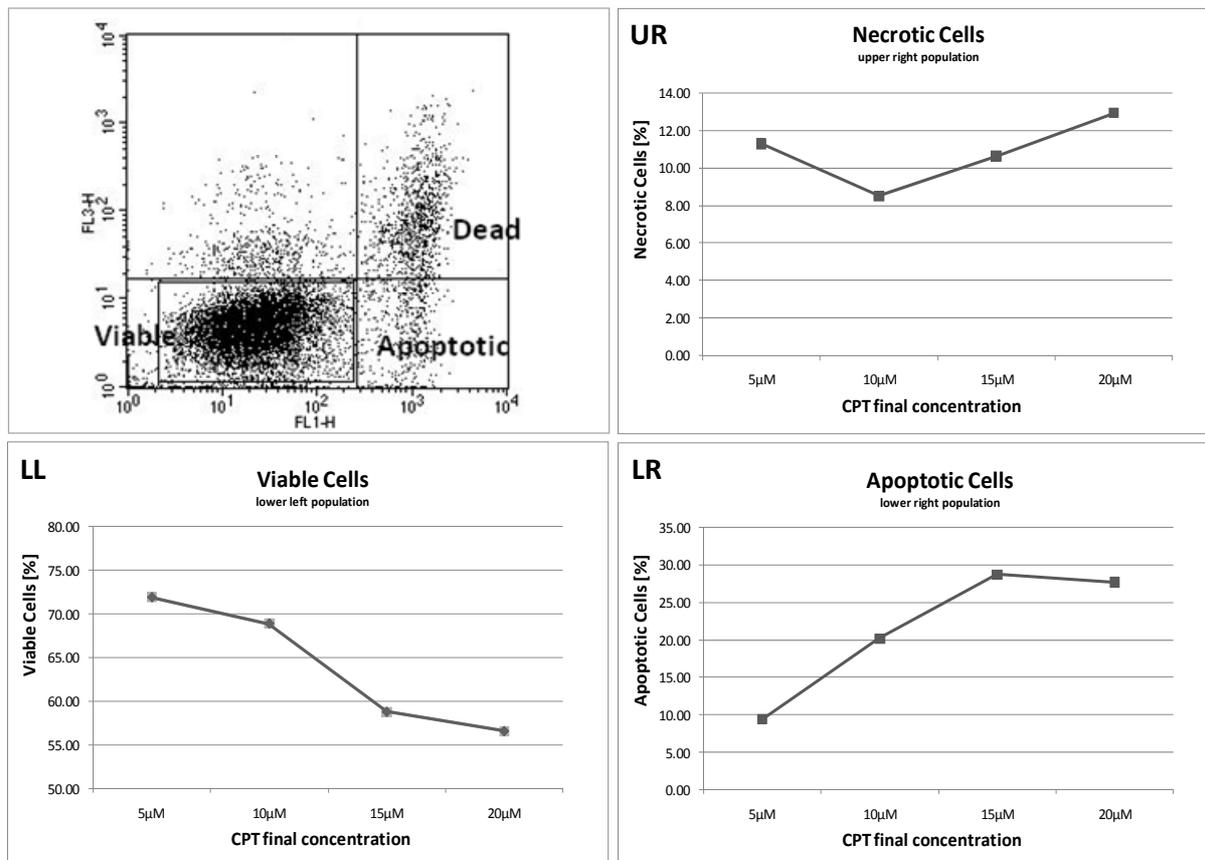
#### 4.3.3.2. Cell Productivity

For run #1 EpoFc titers are plotted in the left graph in Figure 31-B. Due to limitation in the number of samples for ELISA analysis, mock transfected samples were not included in product titer determination. Interestingly, EpoFc-titers steadily increase until 72 hours, followed by a rapid increase between 72 and 96 hours, which is observed for all treated as well as untreated samples. Cumulative viable cell days were determined for each sample and subsequently used for calculation of specific productivity  $q_p$  in secreted attomol per cell and per day (amol/cell/day). As Fig. 31 shows, no effect of transient miR-21 transfection on  $q_p$  was observed, which on the one hand shows that miR-21 knockdown or overexpression does not regulate protein folding or secretion, but also shows that elevated miRNA processing activity within CHO cells does not affect its ability to express recombinant proteins. Since productivity did not seem to be affected, run #2 was performed without determining EpoFc titers.

#### 4.3.4. Stress Resistance of miR-21 Transfected Cells

##### 4.3.4.1. Camptothecin Treatment

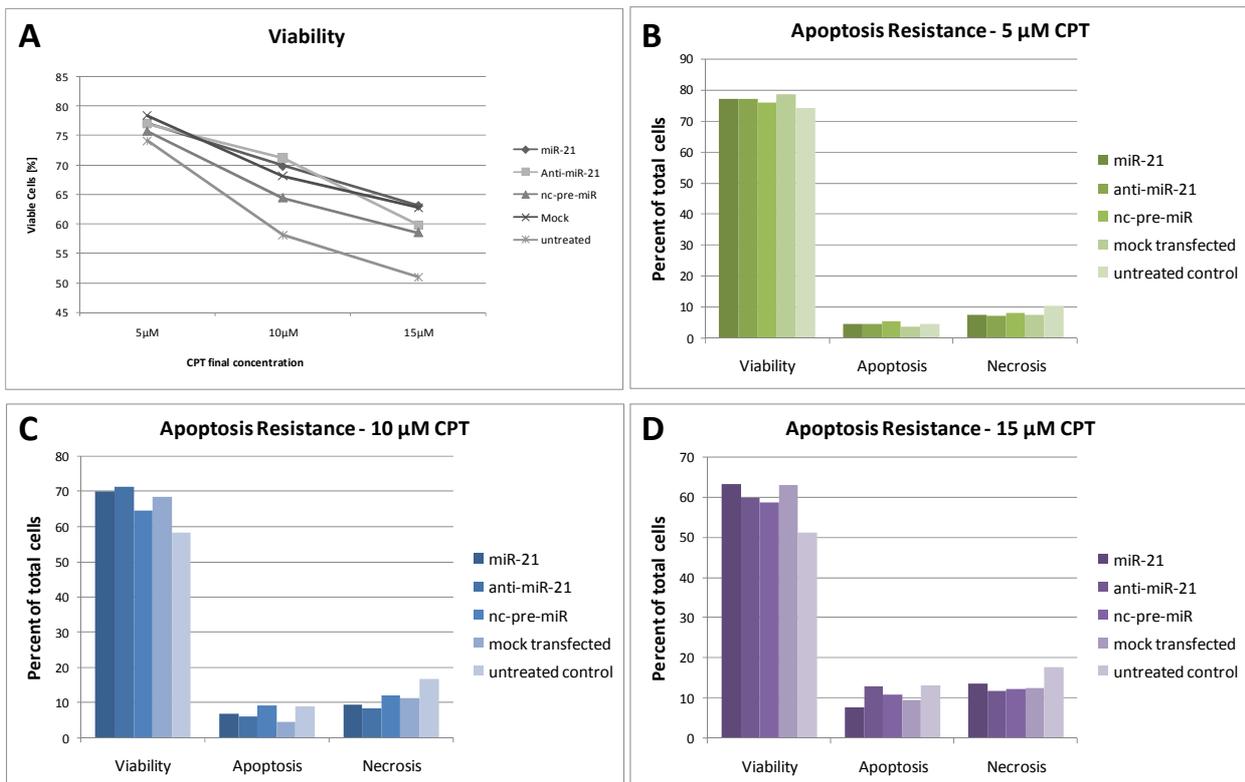
Camptothecin (CPT) is a cancer drug that causes DNA damage by irreversibly inhibition DNA Topoisomerase I; in this study it was used to induce cell death in CHO-EpoFc cells. Figure 32 below shows the results of a preliminary experiment exposing CHO cells to CPT concentrations between 5  $\mu$ M and 20  $\mu$ M for 8 hours. Graphs in Figure 32 are orientated to match the three populations in a FL-1 / FL-3 dot plot, lower left (LL), lower right (LR) and upper right (UR). The LL quadrant corresponds to viable cells (both PI and Annexin-V negative) that decrease steadily with increasing concentrations of CPT. The LR quadrant shows apoptotic cells (Annexin-V positive, PI negative) which increase with higher amounts of CPT. The UR population which depicts dead cells (both, Annexin-V and PI positive) shows marginal increases with higher CPT concentrations. Thus, CPT seems to cause cell death via apoptosis in CHO cells.



**Figure 32 – Preliminary experiment to test concentration dependent effect of CPT on CHO cells.** The graph in the upper left corner shows a typical dot plot of flow cytometry results for apoptosis / cell death analysis, and is divided into 4 quadrants. The graphs labeled with LL, LR and UR show the course of cell percentage for each quadrant over increasing concentrations of camptothecin. While the amount of viable cells decreases and the amount of apoptotic cells increases, the percentage of dead cells remains relatively constant over growing concentrations of CPT.

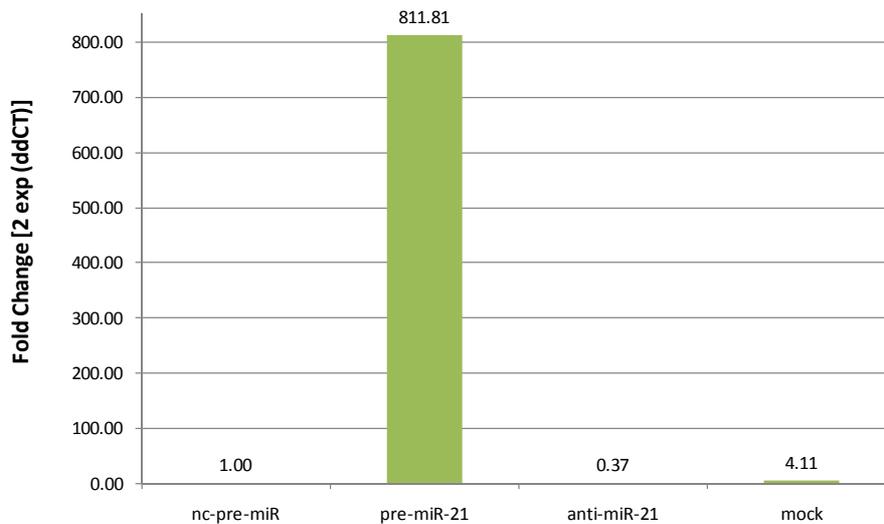
The next experiment was to expose cells with elevated or inhibited endogenous miR-21 levels to CPT and to assess apoptotic resistance: cells were transiently transfected with miR-21 precursors and antagomirs as well as appropriate controls, incubated for 48 hours and finally treated with CPT for 8 hours. Figure 33 summarizes the results obtained for CPT treatment of transfected CHO cells, while Figure 34 confirms high induction as well as knockdown of mature miR-21 levels at the time of treatment. First, compared to the preliminary experiment, less apoptosis was induced this time, with a maximum of 14% for untreated control at 15 μM CPT while seemed to have increased compared to the preliminary experiment. While for 10 μM and especially 15 μM some differences between miR-21 and anti-miR-21 treated cells exist (see bar

chart in Figure 33), these differences cannot be considered as significant. The only constantly observed effect is that untreated cells undergo highest rate of apoptosis and necrosis. Presumably this is due to the fact that untreated cells show fastest growth and therefore their DNA replication machinery might be more active; consequently, since CPT as DNA TOPO I inhibitor has greater impact on more rapidly dividing cells, the untreated control is more prone to camptothecin.



**Figure 33 – Camptothecin treatment of transiently transfected cells:** response of differently transfected cells to CPT treatment using 5, 10 and 15 μM final concentrations. A: course of CHO viabilities of miR-21 transfected cells plus controls over increasing camptothecin (CPT) concentrations. B,C,D: Bar charts depicting cell viability, apoptosis and necrosis for 5 μM, 10 μM and 15 μM CPT concentrations, respectively.

However, there is no significant difference in the response of miR-21 overexpressing or underexpressing cells and nc-pre-miR transfected cells to the induction of cell death using camptothecin. Thus, no relationship between miR-21 overexpression or inhibition and increased apoptosis resistance was detected under these conditions.



**Figure 34 – MicroRNA-21 levels in cells used for CPT treatment.**

#### 4.3.4.2. Nutrient Depletion

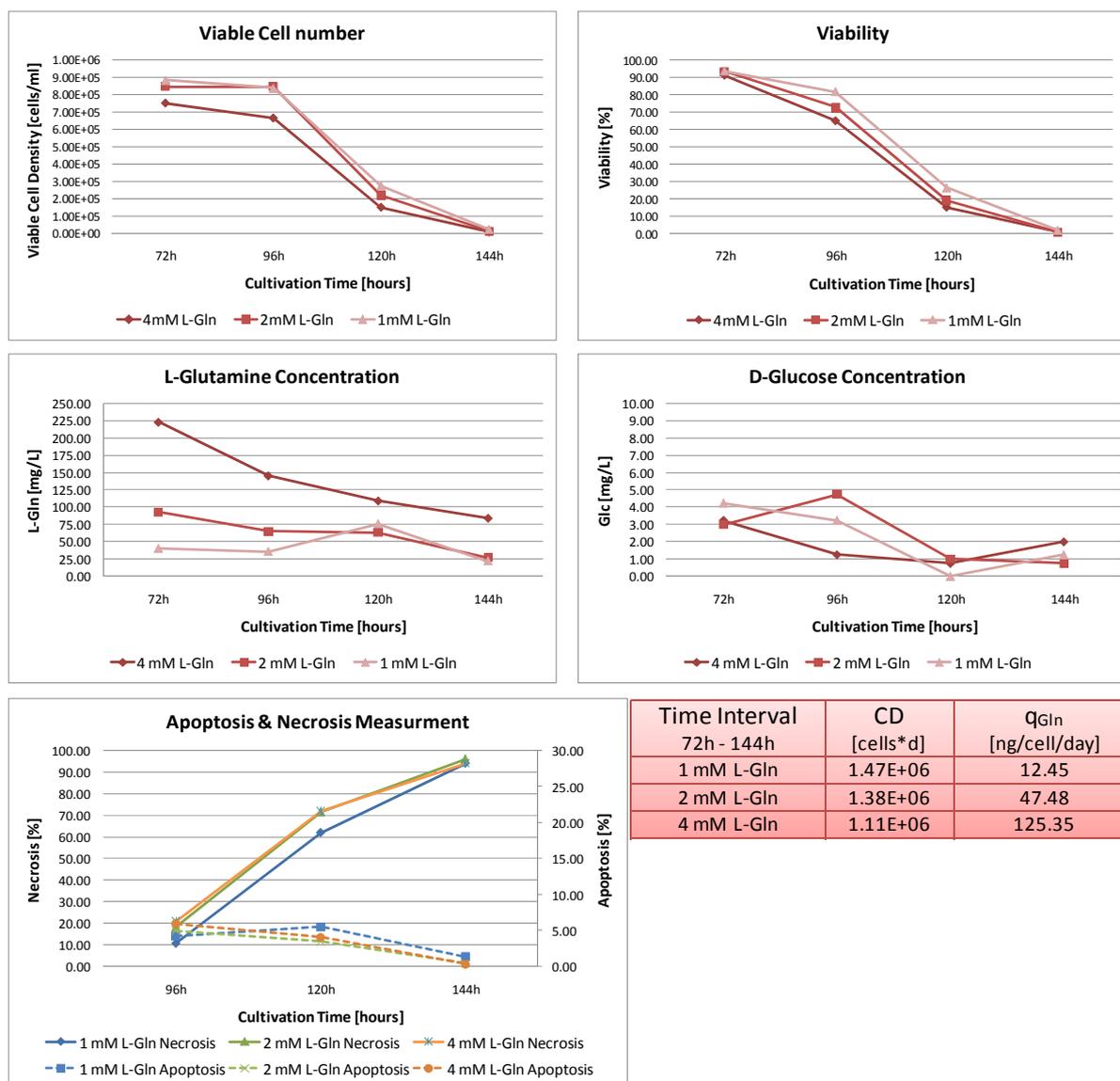
A preliminary experiment was performed in order to determine how CHO-EpoFc cells react to nutrient depletion. Also it was assessed whether a combined limitation of glucose and glutamine had different effects on CHO Epo-Fc culture parameters (i.e. cell growth, viability and apoptosis) than glucose limitation alone. Three media with a decreased glucose concentration of 1 g/L and varying L-glutamin concentrations were tested: (i) medium 1 containing the standard concentration of 4 mM glutamine, medium 2 with 2 mM glutamine, and medium 3 containing only 1 mM glutamine. Figure 35 below shows growth curves and metabolite concentrations for batch cultivations using Media 1 – 3.

As can be seen in graphs showing glucose concentration, the entire amount of 1 g/L glucose has been consumed after 72 hours of cultivation. Consequently, viability and viable cell concentrations start to decrease slowly between 72h and 96h. Cells grown at standard concentration of 4 mM glutamine seem to enter the limitation earlier because their viable cell density (upper left graph Fig. 35) is constantly lower compared to cells grown at 2 mM or 1 mM glutamine and also starts to decrease earlier (after 72h). The limitation in glutamine does not seem to affect the cells in terms of viability since they adapt their specific glutamine consumption rate ( $q_{Gln}$  in ng/cell/day), to the available amount of glutamine. The specific glutamine consumption  $q_{Gln}$  is listed in the lower right Table of Fig. 35, and shows that high L-

gln concentrations result in an increased uptake (4-fold increase in concentration results in a 10-fold faster uptake). However, increased amounts of available glutamine did not result in sustained viability or even higher cell densities but rather seemed to slightly accelerate the cultivation, characterized by an earlier onset of decreasing cell density and viability. This shows that the limit in glucose is the crucial factor for the initiation of cell death due to starvation.

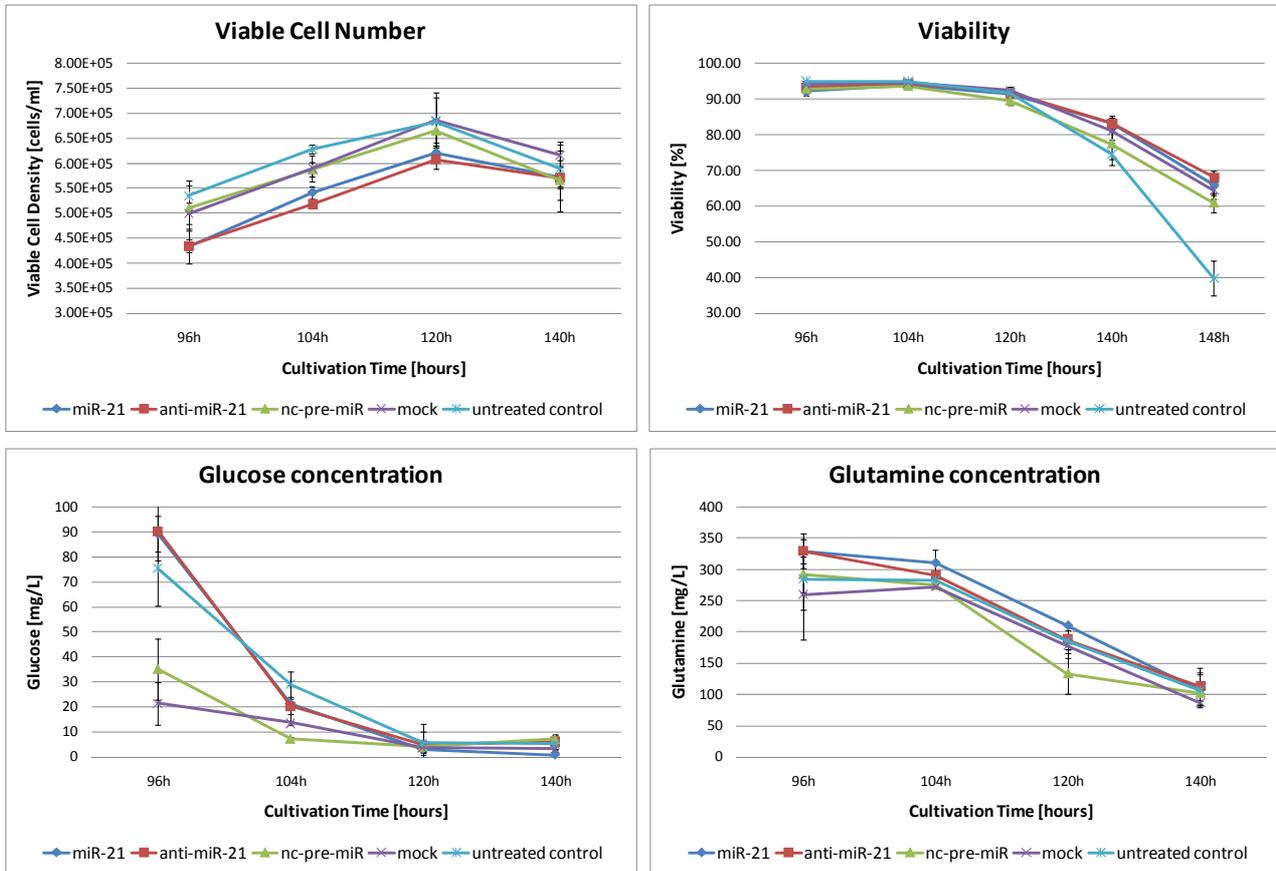
Although it has been shown before that CHO cells enter cell death via apoptosis under nutrient limitation (Hwang and Lee 2008; Hwang and Lee 2008) this effect could not be observed in this study. On the contrary, as the lower left graph in Figure 35 shows, only 15% to 20% of cells are in the state of apoptosis at crucial time points of 96h and 120h. Especially at 96 hours apoptosis would be expected to occur, since then glucose limitation starts to affect viability. Hence, if cells would die via apoptosis they would most likely have to go through programmed cell death at around 96h, so that the high amount of dead cells at 120 hours could be the result of apoptosis. Apparently this is not the case. This could either be because the chosen approach to detect apoptosis via PS (phosphatidylserine) staining with Annexin-V-Fluos was not suitable, or that cells somehow bypassed programmed cell death and died via necrosis, or autophagy.

On account of these preliminary results, medium 1 (1 g/L glucose, 4 mM L-glutamin) was selected for cultivation of transiently transfected cells. Also, since apoptosis did not seem to occur in CHO cells under these conditions, only viability using propidium iodide stain was tested but not apoptosis. This means the small percentages of apoptotic cells were now included in viable cell percentages. Figure 36 shows the courses of viable cell density, viability, glucose and glutamine levels over 140 hours cultivation time subsequent to transfection (148 hours for viability). The graph in Figure 37 depicts the results from Taqman PCR analysis of miR-21 levels after 96h, 120h and 140h.



**Figure 35 – Preliminary experiment for developing a nutrient depletion assay:** The upper graphs show the trend for viable cell density and viability over time. Nutrient limitation was entered between 72h and 96h after cells had reached a density of  $9 \times 10^5$  cells/ml. From this time-point on cell viability and viable cell density decreased steadily. The lower left graph shows that the percentage of dead cells was continuously increasing after the 96 hour time-point. However, only a small percentage of cells entered apoptosis. No significant differences in terms of growth, viability and cell death could be detected between different glutamine concentrations. The center line plots depict glutamine levels on the left and glucose levels on the right. Glucose has been completely consumed in all 3 cultivations at 72 hours. The table in the lower right shows the glutamine consumption rate for CHO cells in different media. Due to the higher availability of glutamine in 4 mM media, cells exhibit higher specific consumption rates. However, increased amounts of available glutamine did not result in increased viability or even higher

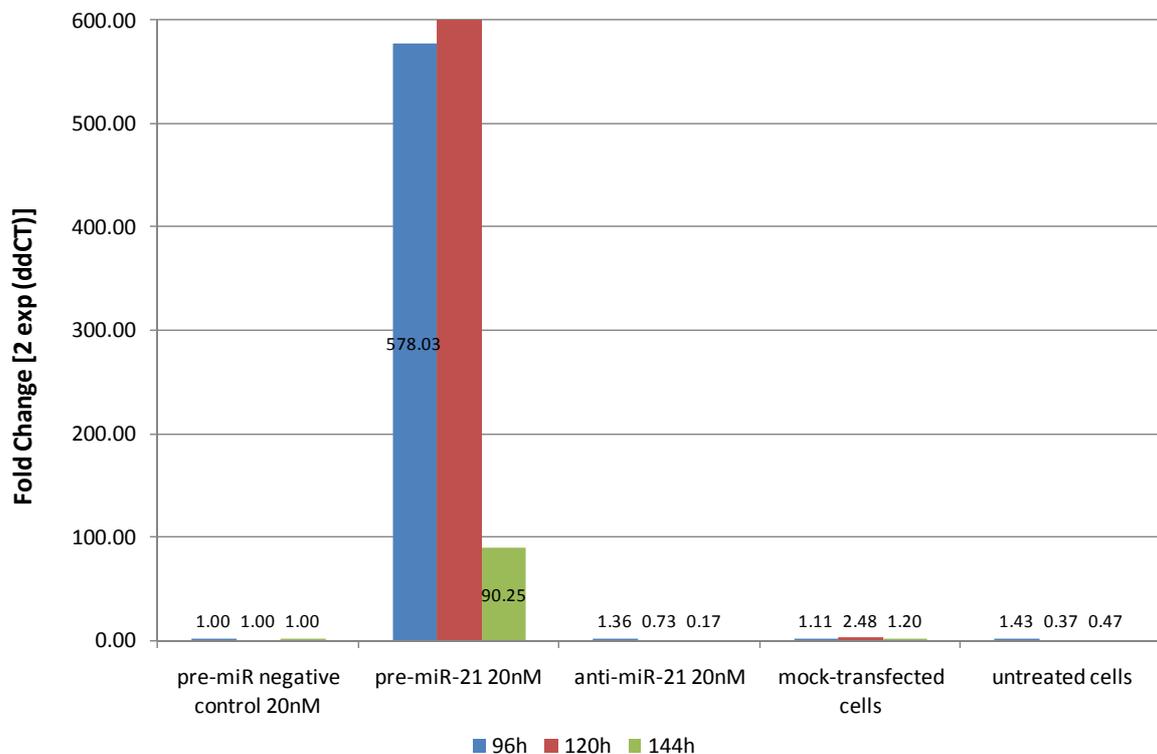
cell densities but seemed to slightly accelerate the cultivation, characterized by an earlier onset of decreasing cell density and viability.



**Figure 36 – Batch cultivation under glucose limitation of transfected cells.** Upper graphs show the course for viable cell density as well as viability; lower graphs depict the course of glucose and glutamine levels in the supernatant.

Cells that were seeded at a concentration of  $1 \cdot 10^5$  cells/ml, reached glucose limitation approximately 104h to 120h after transfection, while glutamine levels were still in the range of 250 – 300 mg/L (approx. 2 mM) at this time. Due to the cells entering limitation, viability started to decrease after 120h, and more rapidly after 140h of batch cultivation. While cells transfected with miR-21 as well as anti-miR-21 exhibited slower growth, lacking approximately  $10^5$  cells/ml behind negative control samples, all transfections showed uniform viability of  $> 90\%$  for 120h after transfections. From this time-point on, untreated control samples exhibited a more rapid decrease in viability, going down to 40% viable cells after 148 hours. This effect was not observed among the other samples, where viabilities remained above 60% until the

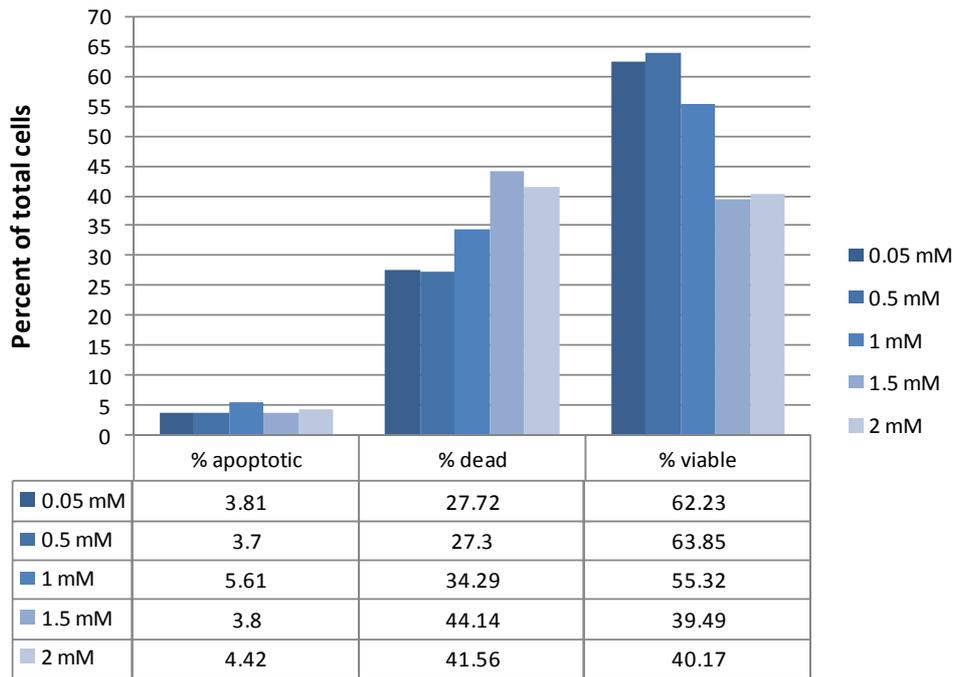
end of the batch culture. Since miR-21 and anti-miR-21 samples exhibited a slower decrease in viability, their viable cell density reached the same level as for the negative control samples at 140h. This effect, even though miR-21 and anti-miR-21 had been growing slower and to lower cell densities, is not due to a later onset of glucose limitation, as can be seen from the glucose graph in the lower left of Figure 36: all samples have entered glucose limitation by hour 104 after transfection. However, a strong effect of miR-21 overexpressing cells leading to sustained viabilities or faster growth was not observed. The general uniformity of viability curves suggests that neither high nor low levels of endogenous miR-21, affect CHO cells under these stress conditions due to limited glucose availability.



**Figure 37 – Confirmation of miRNA induction as well as knockdown for nutrient limitation assay**

#### 4.3.4.3. Oxidative Stress – Hydrogen Peroxide Treatment

Prior to hydrogen peroxide treatment of microRNA transfected cells, a preliminary experiment was performed to determine the effect of H<sub>2</sub>O<sub>2</sub> on CHO cells in terms of concentration dependency and apoptosis induction. Figure 38 below shows the result of a 45 minute treatment with concentrations ranging between 0.05 mM and 2 mM.

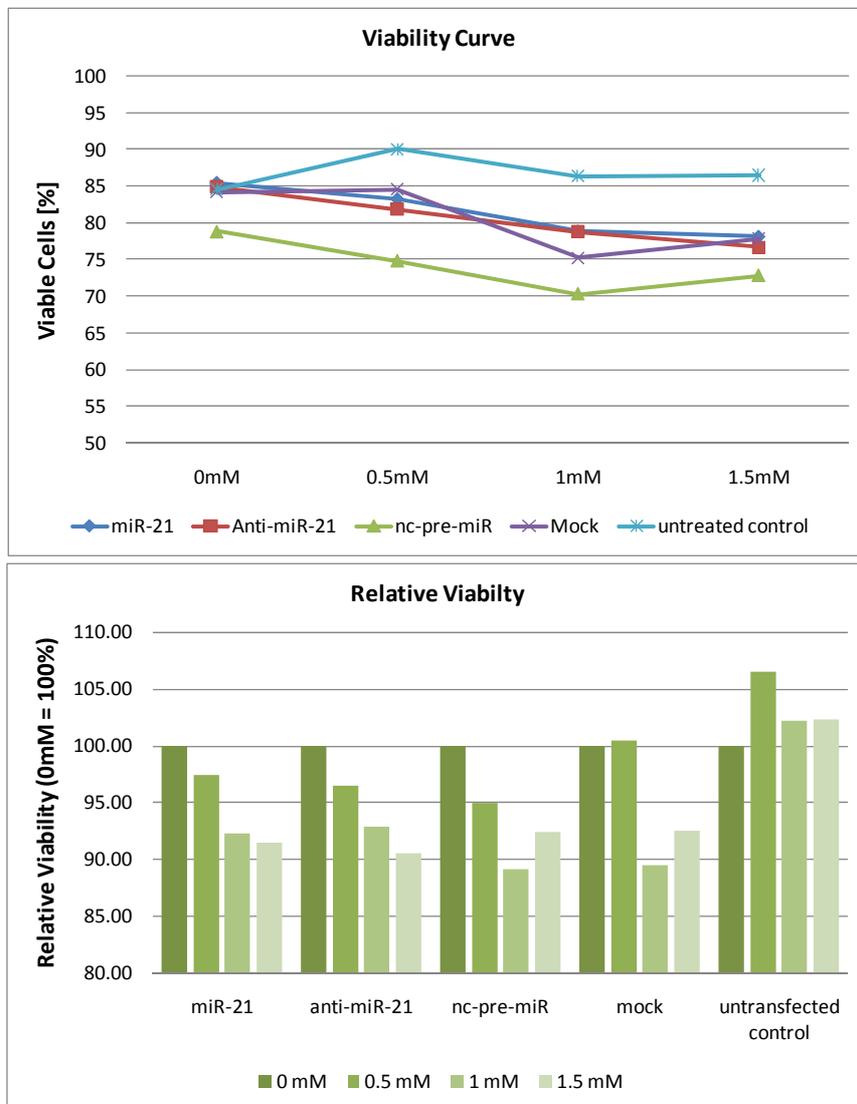


**Figure 38 – Results of a preliminary experiment for oxidative stress induction in CHO cells.**

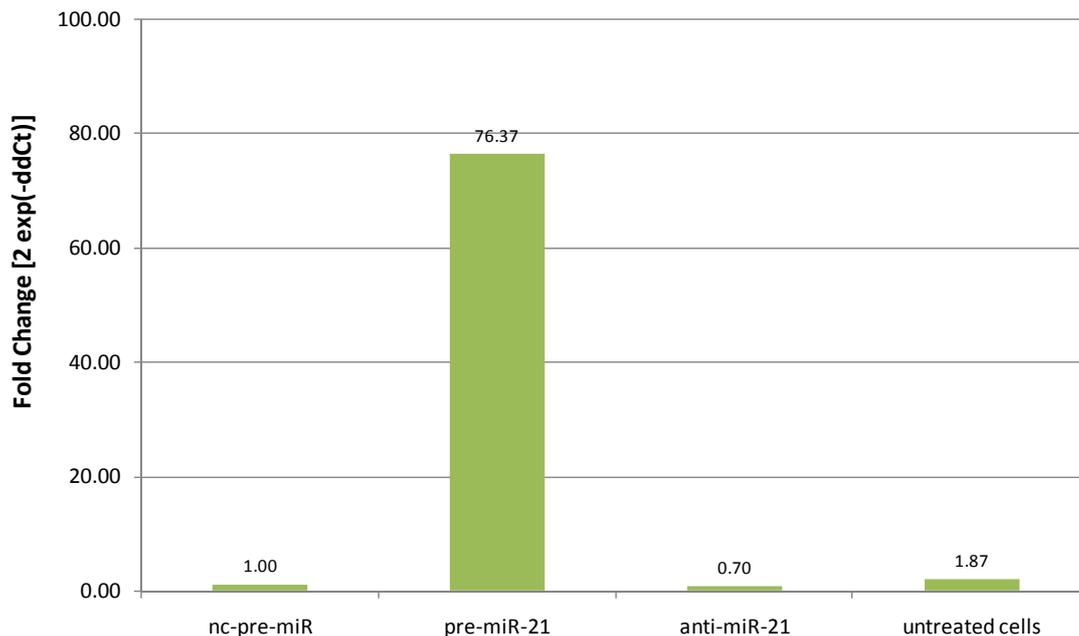
A concentration dependent effect of H<sub>2</sub>O<sub>2</sub> on CHO cells was observed between 0.5 mM and 1.5 mM concentrations and is depicted in Figure 38 as steady increase of dead cells or decrease of viable cell, respectively. Although the percentage of dead cells increased from 27% to 44%, cells did not seem to die via apoptosis as the low levels of Annexin-V positive cells (< 5%) indicate. Even at very low concentrations of 0.05 mM (0 mM control was unfortunately not included) a high percentage of dead cells was detected (27%).

Following the preliminary experiment transiently transfected cells were treated with 3 different H<sub>2</sub>O<sub>2</sub> concentrations and the amount of dead cells was measured. Similar to the preliminary experiment 0 mM H<sub>2</sub>O<sub>2</sub> controls for all transfections exhibit low viabilities of approximately 80% - 85%, shown in the line plot from Fig. 39. The decrease in cell viability over increasing H<sub>2</sub>O<sub>2</sub> concentrations is relatively uniform, especially for miR-21, anti-miR-21 and nc-pre-miR samples, while viabilities of untransfected control samples exhibit no decrease in viability at all. In order to compensate for different starting viabilities at 0 mM H<sub>2</sub>O<sub>2</sub> the bar chart in Fig. 39 shows the relative decrease of viabilities compared to 0 mM H<sub>2</sub>O<sub>2</sub> controls. Obviously neither miR-21

overexpression nor knockdown in CHO cells affects oxidative stress response under these conditions.



**Figure 39 – Effect of hydrogen peroxide treatment on miR-21 transfected cells.** The upper graph shows the course of viability for each transfection over increasing concentrations of hydrogen peroxide. Due to differences in viabilities of 0 mM controls, the lower graph shows the results as a bar chart with 0 mM cell viabilities set to 100%. Here the concentration dependent decrease of viability for miR-21, anti-miR-21 and nc-pre-miR is visible, while mock transfected and especially untransfected cell viabilities did not correlate to increasing H<sub>2</sub>O<sub>2</sub> concentrations.



**Figure 40 – Mature miR-21 levels 48h after transfection (prior to H<sub>2</sub>O<sub>2</sub> treatment).**

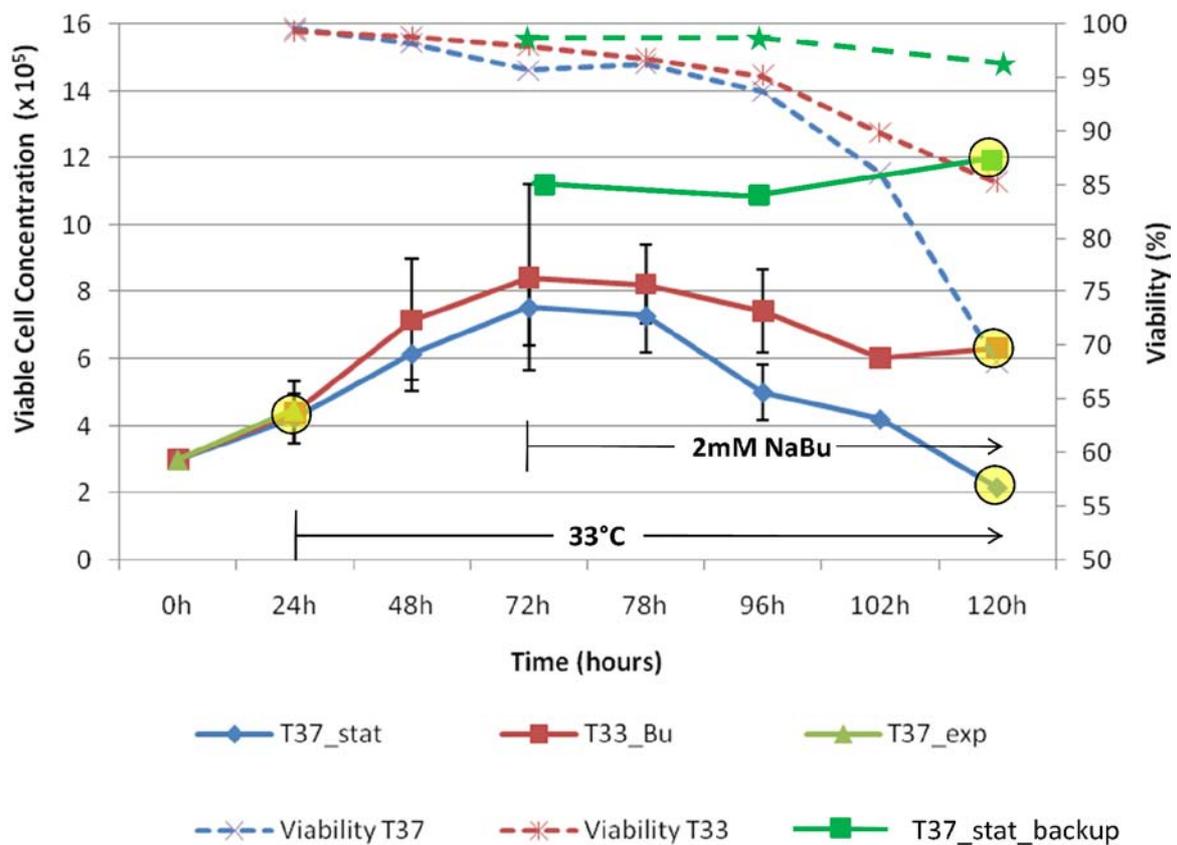
#### **4.4. MicroRNA Profiling Experiments**

For the approach to overexpress or knockdown mature microRNA 21 in CHO cells had not shown promising results under the tested conditions, the identification of other candidate microRNAs for CHO cell line engineering is the logical consequence. Hence, profiling technologies such as microarray studies were planned for the identification of microRNAs expressed CHO as well as potential candidates for overexpression and knockdown studies. The latter are identified by analysis of differentially expressed microRNAs from CHO samples derived from various bioreactor relevant conditions.

##### **4.4.1. Temperature Shift and Sodium Butyrate Treatment**

As previously described (1.2.2), the reduction of cultivation temperature can induce a state of slow growth and high productivity in CHO cells. Temperature shift combined with SME (small molecule enhancers) treatment such as sodium butyrate can further enhance specific productivities and is frequently applied in industrial processes.

In this study we treated CHO-Aven cells with 2 mM NaBu 40 hours after the temperature had been shifted to 33°C (T33\_Bu samples). Total RNA samples were harvested and compared to untreated controls from stationary as well as 'exponential' growth phase (T37\_stat and T37\_exp). Figure 41 below depicts growth and viability of the conducted cultivation. Full lines describe viable cell concentration while the dotted lines show the course of cell viability over time.



**Figure 41 – Growth characteristics of temperature shifted and sodium butyrate treated cells.** *T37\_stat*...cells grown at 37°C until stationary phase was reached at 120h; *T37\_exp*...cells grown at 37°C for 24h; *T33\_Bu*...cells grown at 37°C for 24h, then shifted to 33°C and treated with Sodiumbutyrate (NaBu) after 72h of total cultivation time; *T37\_stat\_backup*...due to problems with the *T37\_stat* culture, these cells were additionally harvested.

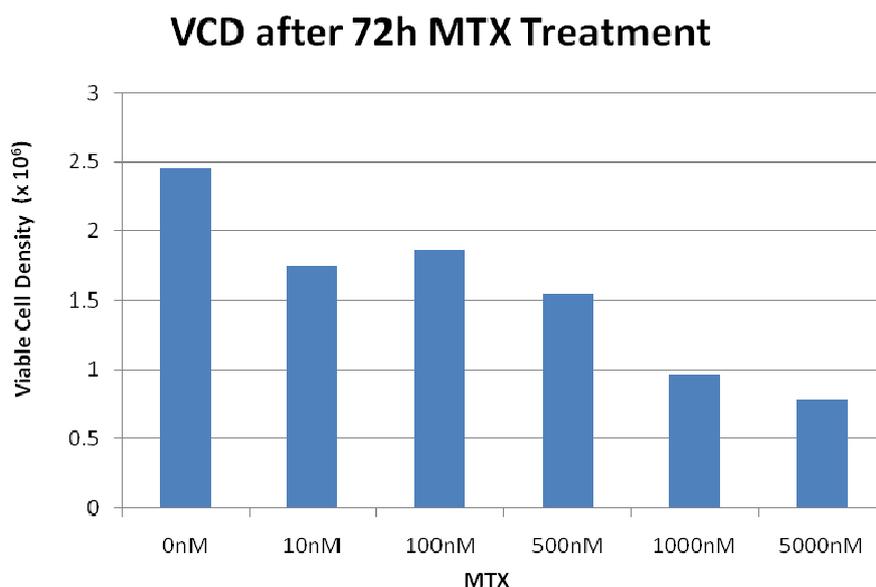
Yellow circles indicate time-points where samples had been taken for total RNA extraction. The red line corresponds to 33°C/NaBu treated cells and is characterized by a lower cell density and a more rapid decrease in viability compared to the 37°C culture (green). The blue curve

corresponds to the original 37°C cultivation, but is not further considered since these cells had been negatively affected probably due to an oxygen limitation following a failure of the stirrer platform between 24 and 48 hours. Instead total RNA samples of a backup culture (green line) were taken. This culture had been started at 0h as well, but not been incorporated in measurements until 72 hour time-point (when problems with T37\_stat become obvious). The backup culture, T37\_stat\_backup, exhibited normal growth and is characterized by higher cell density (up to  $1.2 \times 10^6$  cells/ml) and sustained viability compared to T33\_Bu samples.

All samples have been harvested and are ready for miRNA MicroArray profiling which will be conducted at the University of Minnesota.

#### 4.4.2. Methotrexate Gene Amplification

A second approach to create samples for miRNA profiling was to harvest RNA samples while exposing CHO-MDJ84 cells to methotrexate (MTX) during the process of gene amplification. In order to save time and material it was decided to perform a single-step amplification using only one, elevated, MTX concentration. In a preliminary experiment several MTX concentrations were tested. Viabilities as well as cell densities were measured 72 hours after MTX treatment and are shown in Figure 42 below.



**Figure 42 - Results of preliminary experiment for MTX treatment:** CHO-MDJ-84 cells were treated with varying concentrations of MTX. Cell number and viability were measured after 72h of treatment.

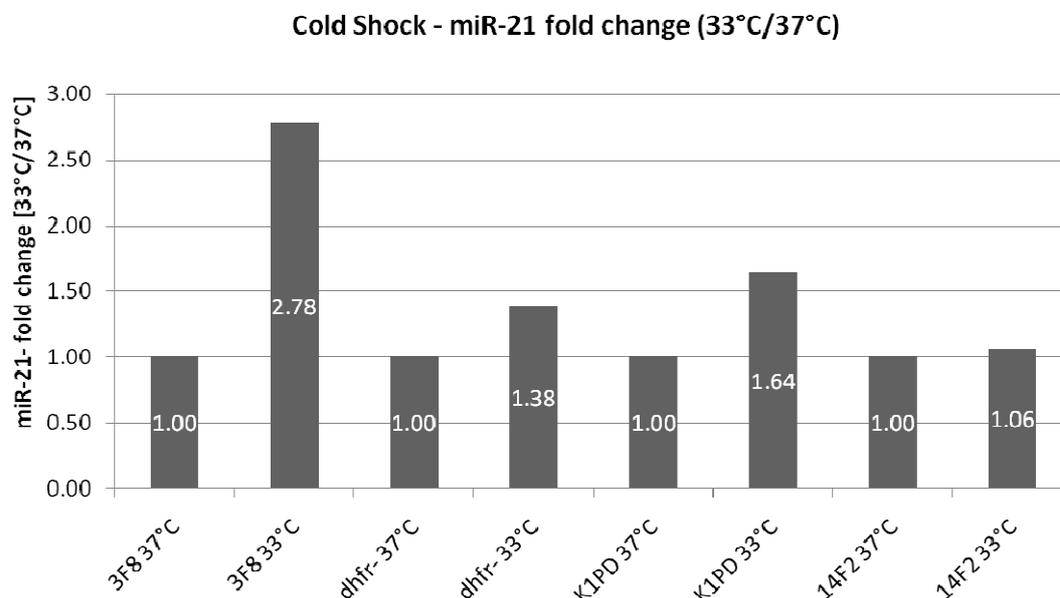
Based on the results obtained from this preliminary experiment, 1 $\mu$ M MTX working concentration was chosen. Samples of MTX treated cells were taken every 5 days and are ready for miRNA microarray profiling and comparison to untreated CHO-MDJ-84 cells. This analysis will be conducted at the University of Minnesota.

#### 4.4.3. Temperature Shift in four CHO Cell Lines – BOKU

Eight total RNA samples, created by Wolfgang Ernst in the course of a project funded by the Austrian Center of Biopharmaceutical Technology (ACBT), extracted from four different cell lines cultivated at both, 33°C and 37°C, were kindly provided for this study.

##### 4.4.3.1. Real-Time qPCR Analysis

Since one aim of this study was to elucidate the role of miR-21 in CHO cells, samples were first subjected to qPCR analysis in order to look for differential regulation of miR-21 under cold-shock conditions; not least because of previous publications reporting upregulation of miR-21 at lower temperature (Gammell, Barron et al. 2007).



**Figure 43 – qPCR detection of changes in miR-21 levels upon a temperature shift from 37°C to 33°C.** In a first step miR-21 levels were normalized to U6-snRNA to calculate  $\Delta\Delta C_t$  values (not shown). Then, for each clone, expression was related to miR-21 levels at 37°C. Four different CHO clones were compared: 3F8 and 14F2 are EpoFc producing clones derived from parental cell line dhfr<sup>-</sup> and K1PD is a CHO clone derived from Sandoz.

The four cell lines used for RNA library preparation included parental cell line dhfr<sup>r</sup>, as well as production clones derived from this cell line (3F8 and 14F2) and a proprietary cell line, K1PD (Sandoz GmbH, Kundl, Austria). For each library qPCR detection of intracellular microRNA-21 levels was performed twice. Figure 43 shows the fold changes of levels in endogenous mature miR-21 (normalized to U6 snRNA expression levels) at 33°C relative to miR-21 levels at 37°C. Two cell lines, K1PD and especially 3F8, exhibit miR-21 upregulation at 33°C. For 14F2 and the dhfr<sup>r</sup> parental cell line we observed no and slight upregulation, respectively.

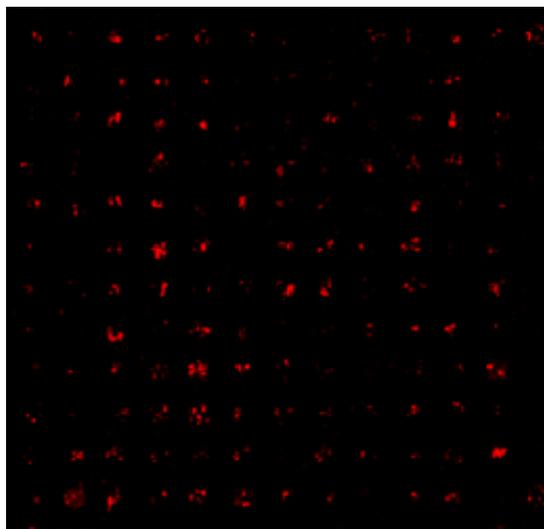
These results have to be considered as preliminary, since further biological replications as well as technical replications of qPCR analysis have to be performed in order to add statistical significance to these results. Nevertheless, the fact that this data is consistent with findings by Gammell et al., suggests a role for miR-21 as regulator (or as being regulated) upon temperature shift.

What remains is to elucidate the causal relationship between the upregulation of a certain set of miRNAs at lower temperature and phenotypic observations such as higher productivity or growth arrest (Kaufmann, Mazur et al. 1999; Trummer, Fauland et al. 2006; Yoon, Ahn et al. 2007; Ahn, Jeon et al. 2008). Recently, Gammell et al. applied 2D Difference Gel Electrophoresis (2D-DIGE) in order to identify differentially expressed proteins upon temperature shift to 31°C. Among the proteins identified as differentially expressed, none of the known targets of miR-21 were down-regulated (Kumar, Gammell et al. 2008).

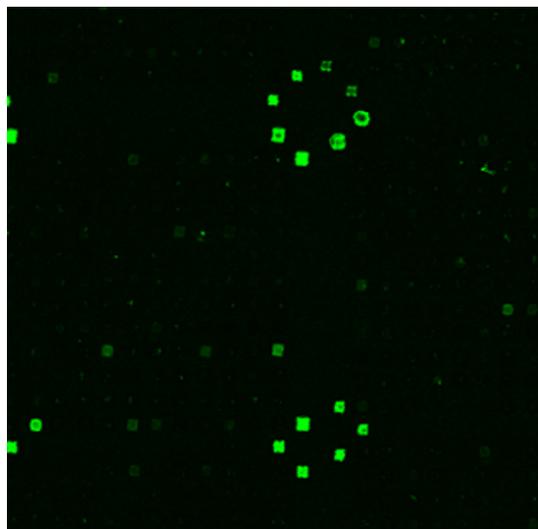
#### 4.4.3.2. Microarray Analysis – microRNA Expression Profiling

Parallel to quantitative PCR analysis total RNA from temperature shift samples was labeled and hybridized to an Exiqon LNA miRNA microarray based on Sanger miRBase version 9.2 (May 2007). Due to problems with dye oxidation (shown in Figure 44), probably caused by elevated ozone levels at the time of measurement, no results for Cy5 labeled samples could be obtained. Consequently, only analyses based on Cy3-labeled 33°C RNA could be performed. However, this allowed presence/absence analysis of miRNAs as well as relative intensity analysis using U6-snRNA-2 spots as reference.

Cy 5  
degradation



Cy 3  
no degradation



**Figure 44 – Cy5 degradation due to dye oxidation:** Screenshot of Cy5 dye degradation resulting in non-evaluable miRNA spots, compared to intact Cy 3 spots.

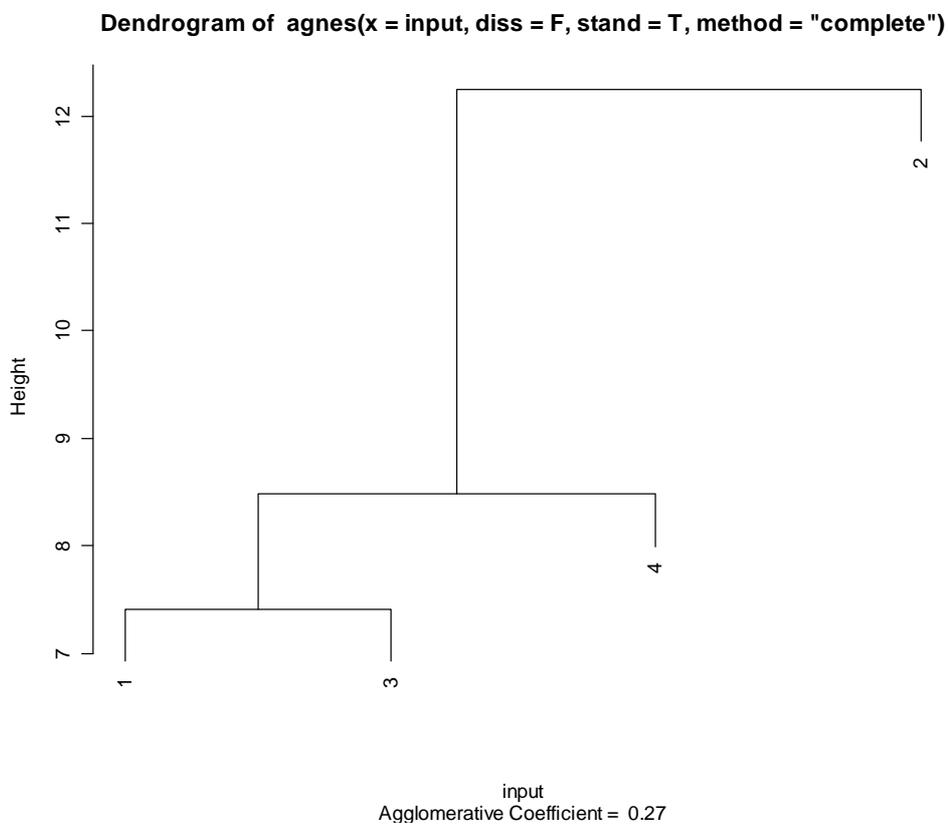
First all Cy3 background corrected spot intensities were analyzed for each chip. Table 14 below summarizes mean, median and quartile values for each library, and gives the amount of high confidence sequences identified. Also, U6-snRNA-2 intensities which were used for signal normalization are given. The summary shows relatively good conformance for intensity distribution (median and quartile values) for all 4 libraries as well as for U6 snRNA intensities. This suggested that an improvised single dye analysis could yield valid results.

A stringent presence/absence analysis of miRNAs yielded a set of 70 mature miRNAs that had exhibited an array signal significantly higher than the background, and a standard deviation across 12 features of 20% or less. Of these, 55 miRNAs were human miRNA sequences, while 12 originate from mouse. Also, three miRPlus sequences were identified which represent proprietary Exiqon sequences that were not yet available in Sanger miRBase at that time. Across all 4 RNA samples (3F8, dhfr, 14F2 and K1PD) 16 miRNAs were commonly present, while 22 miRNAs were present in at least 3 libraries. All miRNAs present on the Exiqon LNA Array are given in Table 15.





lowest level indicating close relationship. Compared to 3F8, K1PD joins this 14F2/dhfr cluster on a significantly lower level, indicating great differences in 3F8 cells compared to the other libraries as far as 33°C miRNA expression is concerned. Taking into account that 3F8 and 14F2 are both derived from dhfr, which would suggest a close relationship, and that K1PD is an independent proprietary cell line provided by Sandoz, these results were not expected and might be due to a bias introduced by the single dye analysis. Results from kmeans clustering confirm these results, when the number of clusters is set arbitrarily set to 3 (not shown).



**Figure 46 - Hierarchical clustering of CHO libraries used for miRNA profiling:** Clustering was performed using R. *Complex linkage* for calculation of distances was chosen. Numbers correspond to libraries according to the labels in Figure 47; 1..14F2, 2..3F8, 3..dhfr, 4..K1PD

## **5. Discussion**

The aim of this study was to gain insights in the roles of microRNAs in CHO cell factories. In the first step Solexa/Illumina deep sequencing technology was used to identify CHO microRNA sequences and to characterize CHO cell lines according to their miRNA expression profile.

### **5.1. MicroRNA Expression Profiling**

We found that CHO microRNAs are characterized by a remarkably high degree of conservation to other mammalian species, mostly human, mouse and rat. Of approximately 12.9 million 22-nt sequences in the Solexa dataset, 11.4 million (88.15%) could be annotated solely based on sequence conservation. Almost the same degree (82%) of conservation was identified for 23-nt sequences. Among these sequences we identified 260 distinct mature miRNA sequences, which exhibited a range of abundances, from only few up to several thousand clonal copies present in one library. Consequently, compared to almost 600 microRNAs (Sanger miRBase release 13.0, March 2009) currently known in human, a single run of next-generation sequencing yielded a considerable amount of miRNA sequences for CHO and again proved to be a powerful tool in functional genomics as has been shown before (Kuchenbauer, Morin et al. 2008; Morin, O'Connor et al. 2008). Both, the degree of conservation as well as the large frequencies of single microRNAs suggest that miRNAs might carry important biological functions in Chinese Hamster Ovary cells. Thus, the elucidation of miRNA regulation in CHO, with special emphasis on pathways and programs important for industrial processes, can yield promising candidates for cell line engineering. Although for now, the lack of genomic sequences of CHO cells hampers the use of sophisticated algorithms to predict and look for novel CHO miRNAs, the availability of 260 CHO microRNA sequences is an important gain of knowledge and will prove very useful.

Solexa abundance informations have been used before to describe miRNA expression levels for sequenced small RNA libraries (Kuchenbauer, Morin et al. 2008; Morin, O'Connor et al. 2008). Similarly, in this study we used normalized Solexa expression levels to point out differences in terms of microRNA expression between CHO parental cell lines, production clones and CHO Aven batch cultivation. We visualized the data in a heat-map and clustered the libraries

according to their miRNA abundance per million reads. What we found was, that large differences in microRNA expression are more likely to occur due to profound genetic changes such as stable transfection and subsequent gene amplification than due to changes in culture conditions (such as serum-free adaption, temperature shift or butyrate treatment). This observation resembles what several studies have already shown, which is that microRNA regulation is highly cell line and tissue specific (Cheng, Byrom et al. 2005; Lim, Lau et al. 2005; Gammell, Barron et al. 2007; Si, Zhu et al. 2007; Zhu, Wu et al. 2008): (i) adverse effects of the same microRNA in different cell lines have been published; several microRNAs (ii) are only expressed in certain cell lines or types of tissues, (iii) are highly expressed or inhibited in cells that have become malignant and have formed tumors, or (iv) are induced upon change of external conditions.

This creates the need of in-depth characterization of microRNA expression for each organism as well as for individual cell lines. On account of numerous publications that brought miR-21 in context with cell proliferation, transformation and cell death, it was decided to assess the effect of overexpression and inhibition of miR-21 in an industrially relevant CHO cell line. This decision was also justified by the fact that miR-21 was (and by now is) the first publicly available microRNA sequence for Chinese Hamster.

## **5.2. Functional Study of microRNA-21 in CHO**

In a first approach the regulatory functions of CHO microRNAs need to be assessed by identifying target mRNAs; subsequently gene ontology can be used to link sets of miRNAs to specific cellular pathways. Müller and Grillari et al. took a first step into this direction by summarizing publicly available miRNA data and linking it with cellular programs relevant for biotech industry, such as cell growth, death or stress resistance (Muller, Katinger et al. 2008). Here we were able to identify a putative target of miR-21 – PDCD4: programmed cell death 4, or neoplastic transformation inhibitor has originally been identified as a mouse messenger RNA up-regulated in several cell types during apoptosis. Thus, it has been granted a role as important tumor suppressor within mammalian cells and, in the recent past, has been frequently described as down-regulated by miR-21 resulting for instance in increased

invasiveness of cells (Asangani, Rasheed et al. 2008; Lu, Liu et al. 2008; Zhu, Wu et al. 2008). Consequently, the PDCD4 3'UTR target site for miR-21 became publicly available and, using repositories such as targetscan, high conservation of this 3'UTR region was detected. Due to the collaboration with the Austrian Center for Biopharmaceutical Technology it was possible to derive CHO EST sequences that had been annotated to human PDCD4. Looking closer at these sequences we identified the presence of the PDCD4:miR-21 8mer binding site within two CHO PDCD4 singlets. In contrast to the high sequence variation often found for non-coding parts of mRNA we observed high sequence identity between CHO singlets and human PDCD4 3'UTR. Also, it seems that the miR-21 binding site is located within an alternative PDCD4 splice site in CHO which gives interesting implications for miR-21 functionality: is it possible that PDCD4 can escape miR-21 regulation by alternative splicing? Sandberg et al. have previously shown that proliferating cells can express mRNA isoforms with shortened 3' UTRs for example due to upstream polyadenylation sites or alternative splicing (Sandberg, Neilson et al. 2008), which do not underlie microRNA regulation. Could inordinately high expression of miR-21 (that turns miR-21 from a proto-oncomir into an oncomir) activate SR-proteins or other factors that regulate splicing to remove its PDCD4 target site so that the effect of miR-21 overexpression is downsized?

Besides postulating speculations like these the logical consequence was to study the effect of miR-21 overexpression as well as inhibition in CHO cell factories. Therefore, several assays were developed in order to assess cell growth, productivity, apoptosis and stress resistance upon transient transfection of microRNAs.

However, batch cultivation of miR-21 and anti-miR-21 transfected cells had no significant impact on cell growth, viability or productivity compared to negative controls in CHO. While in the case of productivity measurable effects of miR-21 would have been somewhat surprising, it was to a certain extent expected that either overexpression or knockdown of miR-21 would affect CHO cell growth; not least because of the putative interaction with CHO PDCD4 which has been shown to act as a tumor suppressor by down-regulating MAP4K1 which results in reduced invasiveness of cells (Yang, Matthews et al. 2006). Consequently miR-21, by repressing

translation of PDCD4, has been shown to increase invasiveness and to be responsible for in vivo growth advantages of tumor cells (Si, Zhu et al. 2007). Also increased invasiveness, intravasation or metastasis have been observed for attachment dependent cancer cell models upon miR-21 overexpression (Asangani, Rasheed et al. 2008). Yet, this study shows that cells adapted to grow independent of surface attachment and which have therefore lost direct contact to each other, do not respond to increased levels of miR-21 as far as cell growth is concerned. Thus, it seems that enhancing invasiveness of CHO cells does not result in increased growth rates of suspension cells.

Based on these results, assays were developed to test the impact of miR-21 overexpression and inhibition on CHO cells upon induction of cellular stress. Camptothecin (CPT) has been shown before to induce apoptosis in mammalian cells via inhibition Topo Isomerase I and cause of DNA damage (Hinz, Helleday et al. 2003; Holme, Yadav et al. 2007). Thus, miR-21 transfected cells were treated with CPT and parameters such as cell growth and viability (including apoptosis) were measured. However, no sustained viabilities under CPT treatment could be observed in transfected cells, relative to negative controls. In order to create stress assays that resemble common problems in bioreactors, nutrient limitation was induced in CHO cells using a glucose deprived DMEM/Ham's F12 medium. However, an experiment consisting of three biological replicates and two technical replicates of each measurement, confirmed that neither miR-21 overexpression nor knockdown could improve cell viabilities under these conditions. Finally, in a similar approach to CPT treatment, the effect of hydrogen peroxide on CHO cells was assessed but did not yield promising results under the chosen conditions either.

### **5.3. Perspectives**

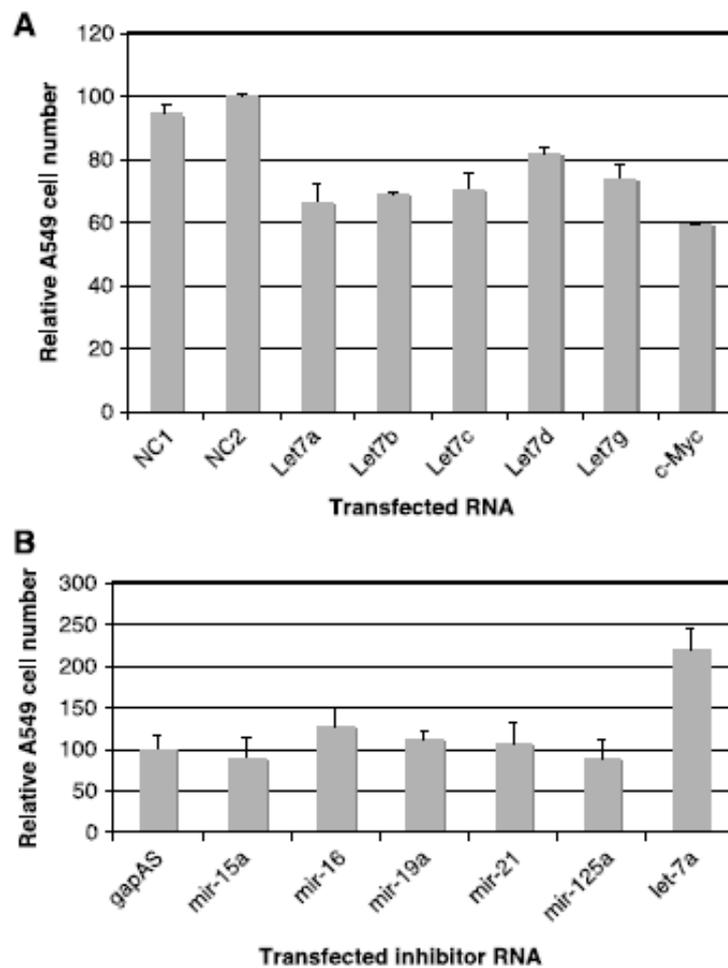
These results created the need to look for additional microRNAs with potential use in cell line engineering. Therefore effort was put into profiling of CHO microRNAs under various bioreactor relevant conditions: for example gene amplification is a very important step for achieving high space-time yields, and is often achieved using methotrexate (MTX) treatment of dhfr<sup>-</sup> CHO cells. Thus, samples for miRNA profiling were for example generated from CHO MDJ84 cells undergoing 1  $\mu$ M MTX amplification. Also biphasic processes are commonly used in

mammalian-cell based bioprocesses, for example via shifting the temperature to 33°C and adding sodium butyrate to the media. RNA from CHO Aven cells undergoing such biphasic cultivation were generated and are ready for analysis at Professor Hu's group at the University of Minnesota. At the University of Applied Life Sciences and Natural Resources in Vienna, CHO RNA samples from temperature shift cultivations were available through the ACBT; samples were labeled and hybridized to human and mouse miRNA arrays but artefacts introduced by dye oxidation made analyses delicate. Eventually it was possible to conduct presence/absence analysis on the dataset and to gain some information about miRNA expression under 33°C.

Therefore it is absolutely necessary to repeat profiling studies on available samples, as well as to conduct sophisticated experiments to create new samples for miRNA differential expression analysis. This is because the better CHO cells are characterized in terms of miRNA expression, the more useful public available data on miRNA regulation can be. Several microRNAs have been shown to regulate cellular programs important for cell engineering (Muller, Katinger et al. 2008), and knowing their about their presence or absence or even their levels of expression is crucial. In the following example we show how information created in this study can be used to generate miRNA-based cell line engineering strategies.

We showed that Solexa sequencing as well as cross-species microarray experiments have identified microRNA let-7a as highly expressed in CHO. Interestingly, in two cancer cell lines, Namalwa cells (resembling Burkett's Lymphoma) and DLD-1 human colon cancer cells, transfection and subsequent overexpression of mature let-7a had resulted in reduced cell growth (Akao, Nakagawa et al. 2006; Sampson, Rong et al. 2007). As shown in Figure 47 below, another study identified let-7a as a global inhibitor of cell growth by repressing cell proliferation pathways in human lung cancer cells (Johnson, Esquela-Kerscher et al. 2007). This has been suggested to be the consequence of let-7a interacting with several cell proliferation proteins, such as ras and c-myc. As a transcription factor c-myc is frequently found over-expressed in cancer cell lines (Sampson, Rong et al. 2007) where it was shown to drive cell proliferation via activation of cdc25 (an enzyme kinase that activates cdk/cyclins) and therefore received the status of a proto-oncogene. Figure 47 below depicts results from let-7a overexpression and

inhibition from Akao and Nakagawa et al. from 2006; as can be seen in B of Fig. 47, the transfection of let-7a inhibitors results in a 100% increase in cell density. These findings and the high levels of let-7a which were identified in CHO in this study force the question what happens if let-7a levels are reduced in CHO using for example anti-let-7a inhibitors? At the same time inhibitor of miR-15 could be introduced in order to avoid onset of apoptosis, because miR-15 as well as miR-16 have been shown to downregulate expression of anti-apoptotic Bcl-2 (Cimmino, Calin et al. 2005). Presence of miR-15b at 33°C was shown in this study, thus there would be space for a miR-15 knockdown to create an increased apoptosis resistance via Bcl-2.



**Figure 47 – Effect of hsa-let-7a in lung cells:** (A) Transfection of let-7 precursor negatively affects cell growth of A549 cells compared to two negative controls (NC1, NC2) and similar to c-Myc siRNA (c-Myc). (B) Transfection of anti-let-7a inhibitors resulted in a 100% increase in cell number relative to negative controls and other miRNA inhibitors. Figure taken from page 7715 from (Akao, Nakagawa et al. 2006).

While an increase in cell growth and viability improves the space-time-yield of a bioprocess, increasing the recombinant protein production rate  $q_p$  is another approach. MicroArray profiling experiments in this study identified microRNA-10b, a close homolog to miR-10a, expressed at very low levels in three out of four libraries. This is of interest because miR-10a takes an exceptional position among miRNAs: it has been shown to bind its target mRNAs in the 5' untranslated regions (5' UTR) resulting in increased expression of target proteins. An effect of miR-10a was for example detected upon amino acid starvation where it was shown to enhance expression of ribosomal proteins (RPs) and consequently increased biogenesis of ribosomes. Also, overexpression of miRNA miR-10a resulted in an overall increase of 30% of global protein production, determined by [<sup>35</sup>S] labeling of newly synthesized proteins (Orom, Nielsen et al. 2008). This has further implications for cancer development due to enhanced oncogenic transformation events observed upon miR-10a overexpression. However, these observations raise the question whether enhanced protein production can be achieved in CHO cells upon miR-10a overexpression. Not least since CHO cells normally exhibit low miR-10 expression, even at low temperatures where energy is shifted from cell division to protein production (Kaufmann, Mazur et al. 1999).

In order to find additional interesting microRNA targets for engineering CHO cells, our current knowledge about microRNA expression and their biological functions in this rodent species needs to be increased. In a first approach to reach this goal we have proven next-generation sequencing technologies such as Solexa/Illumina as well as microRNA profiling experiments using cDNA microarray technology to be powerful tools and to be of considerable use in this endeavor. In the next step, to increase the number of known CHO microRNA sequences, additional small RNA libraries need to be created (ideally including a variety of CHO tissues) and sequenced. In parallel the search for CHO specific microRNAs using bioinformatic tools has to be continued. Subsequently this sequence knowledge can be used for the development of CHO specific microRNA microarrays which will be more powerful in profiling miRNA expression and identifying differentially expressed genes. In the end results from such studies will allow us to understand the function of certain miRNA well enough so that the first miRNA engineered host cell line for recombinant protein production can be created.

## 6. References

- Ahn, W. S., J. J. Jeon, et al. (2008). "Effect of culture temperature on erythropoietin production and glycosylation in a perfusion culture of recombinant CHO cells." *Biotechnol Bioeng* 101(6): 1234-44.
- Akao, Y., Y. Nakagawa, et al. (2006). "let-7 microRNA functions as a potential growth suppressor in human colon cancer cells." *Biol Pharm Bull* 29(5): 903-6.
- Alberts, J., Lewis, Raff, Roberts, Walter (2002). *Molecular biology of the cell*. New York, NY 10016, Garland Science.
- Ambros, V. (2004). "The functions of animal microRNAs." *Nature* 431(7006): 350-5.
- Antoniou, M., L. Harland, et al. (2003). "Transgenes encompassing dual-promoter CpG islands from the human TBP and HNRPA2B1 loci are resistant to heterochromatin-mediated silencing." *Genomics* 82(3): 269-79.
- Arden, N. and M. J. Betenbaugh (2004). "Life and death in mammalian cell culture: strategies for apoptosis inhibition." *Trends Biotechnol* 22(4): 174-80.
- Asangani, I. A., S. A. Rasheed, et al. (2008). "MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pcd4 and stimulates invasion, intravasation and metastasis in colorectal cancer." *Oncogene* 27(15): 2128-36.
- Bartel, D. P. (2004). "MicroRNAs: genomics, biogenesis, mechanism, and function." *Cell* 116(2): 281-97.
- Bentley, D. R., S. Balasubramanian, et al. (2008). "Accurate whole human genome sequencing using reversible terminator chemistry." *Nature* 456(7218): 53-9.
- Berezikov, E., V. Guryev, et al. (2005). "Phylogenetic shadowing and computational identification of human microRNA genes." *Cell* 120(1): 21-4.
- Brennecke, J., A. Stark, et al. (2005). "Principles of microRNA-target recognition." *PLoS Biol* 3(3): e85.
- Chan, J. A., A. M. Krichevsky, et al. (2005). "MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells." *Cancer Res* 65(14): 6029-33.
- Chao, D. T., G. P. Linette, et al. (1995). "Bcl-XL and Bcl-2 repress a common pathway of cell death." *J Exp Med* 182(3): 821-8.
- Chen, C., D. A. Ridzon, et al. (2005). "Real-time quantification of microRNAs by stem-loop RT-PCR." *Nucleic Acids Res* 33(20): e179.

Chen, Y., W. Liu, et al. (2008). "MicroRNA-21 down-regulates the expression of tumor suppressor PDCD4 in human glioblastoma cell T98G." Cancer Lett 272(2): 197-205.

Cheng, A. M., M. W. Byrom, et al. (2005). "Antisense inhibition of human miRNAs and indications for an involvement of miRNA in cell growth and apoptosis." Nucleic Acids Res 33(4): 1290-7.

Cimmino, A., G. A. Calin, et al. (2005). "miR-15 and miR-16 induce apoptosis by targeting BCL2." Proc Natl Acad Sci U S A 102(39): 13944-9.

Cmarik, J. L., H. Min, et al. (1999). "Differentially expressed protein Pcd4 inhibits tumor promoter-induced neoplastic transformation." Proc Natl Acad Sci U S A 96(24): 14037-42.

Das, P. P., M. P. Bagijn, et al. (2008). "Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline." Mol Cell 31(1): 79-90.

Dinnis, D. M. and D. C. James (2005). "Engineering mammalian cell factories for improved recombinant monoclonal antibody production: lessons from nature?" Biotechnol Bioeng 91(2): 180-9.

Dunster, C. A., K. H. Cheeseman, et al. (1997). "The effect of oxidative stress on the production of the recombinant protein, interferon gamma, produced by Chinese hamster ovary cells in stirred-batch culture." Appl Microbiol Biotechnol 48(2): 198-203.

Durrschmid, M. P., K. Landauer, et al. (2003). "Comparison of fluidized bed and ultrasonic cell-retention systems for high cell density mammalian cell culture." Biotechnol Prog 19(3): 1045-8.

Figueroa, B., Jr., E. Ailor, et al. (2007). "Enhanced cell culture performance using inducible anti-apoptotic genes E1B-19K and Aven in the production of a monoclonal antibody with Chinese hamster ovary cells." Biotechnol Bioeng 97(4): 877-92.

Figueroa, B., Jr., T. M. Sauerwald, et al. (2001). "Comparison of Bcl-2 to a Bcl-2 deletion mutant for mammalian cells exposed to culture insults." Biotechnol Bioeng 73(3): 211-22.

Figueroa, B., Jr., T. M. Sauerwald, et al. (2003). "A comparison of the properties of a Bcl-xL variant to the wild-type anti-apoptosis inhibitor in mammalian cell cultures." Metab Eng 5(4): 230-45.

Filipowicz, W., S. N. Bhattacharyya, et al. (2008). "Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?" Nat Rev Genet 9(2): 102-14.

Frankel, L. B., N. R. Christoffersen, et al. (2008). "Programmed cell death 4 (PDCD4) is an important functional target of the microRNA miR-21 in breast cancer cells." J Biol Chem 283(2): 1026-33.

Gammell, P., N. Barron, et al. (2007). "Initial identification of low temperature and culture stage induction of miRNA expression in suspension CHO-K1 cells." J Biotechnol 130(3): 213-8.

Gandor, C., C. Leist, et al. (1995). "Amplification and expression of recombinant genes in serum-independent Chinese hamster ovary cells." FEBS Lett 377(3): 290-4.

Girod, P. A., M. Zahn-Zabal, et al. (2005). "Use of the chicken lysozyme 5' matrix attachment region to generate high producer CHO cell lines." Biotechnol Bioeng 91(1): 1-11.

Goswami, J., A. J. Sinskey, et al. (1999). "Apoptosis in batch cultures of Chinese hamster ovary cells." Biotechnol Bioeng 62(6): 632-40.

Greber, D. and M. Fussenegger (2007). "Multi-gene engineering: simultaneous expression and knockdown of six genes off a single platform." Biotechnol Bioeng 96(5): 821-34.

Grimson, A., K. K. Farh, et al. (2007). "MicroRNA targeting specificity in mammals: determinants beyond seed pairing." Mol Cell 27(1): 91-105.

Hammond, S. M. (2006). "MicroRNAs as oncogenes." Curr Opin Genet Dev 16(1): 4-9.

Hammond, S. M. (2006). "RNAi, microRNAs, and human disease." Cancer Chemother Pharmacol 58 Suppl 1: s63-8.

Harris, M. H. and C. B. Thompson (2000). "The role of the Bcl-2 family in the regulation of outer mitochondrial membrane permeability." Cell Death Differ 7(12): 1182-91.

Hinz, J. M., T. Helleday, et al. (2003). "Reduced apoptotic response to camptothecin in CHO cells deficient in XRCC3." Carcinogenesis 24(2): 249-53.

Holme, A. L., S. K. Yadav, et al. (2007). "Automated laser scanning cytometry: a powerful tool for multi-parameter analysis of drug-induced apoptosis." Cytometry A 71(2): 80-6.

Houbaviy, H. B., L. Dennis, et al. (2005). "Characterization of a highly variable eutherian microRNA gene." RNA 11(8): 1245-57.

Hwang, S. O. and G. M. Lee (2008). "Autophagy and apoptosis in Chinese hamster ovary cell culture." Autophagy 4(1): 70-2.

Hwang, S. O. and G. M. Lee (2008). "Nutrient deprivation induces autophagy as well as apoptosis in Chinese hamster ovary cell culture." Biotechnol Bioeng 99(3): 678-85.

Jayapal, K. F. W., Miranda G. S. Yap, Wei-Shou Hu (2005). Recombinant Protein Therapeutics from CHO Cells - 20 Years and Counting. CHO Consortium - SBE Special Section, Society for Biological Engineering: 8.

Jiang, H. and W. H. Wong (2008). "SeqMap: mapping massive amount of oligonucleotides to the genome." Bioinformatics 24(20): 2395-6.

- John, B., A. J. Enright, et al. (2004). "Human MicroRNA targets." PLoS Biol 2(11): e363.
- Johnson, C. D., A. Esquela-Kerscher, et al. (2007). "The let-7 microRNA represses cell proliferation pathways in human cells." Cancer Res 67(16): 7713-22.
- Jones-Rhoades, M. W., D. P. Bartel, et al. (2006). "MicroRNAs and their regulatory roles in plants." Annu Rev Plant Biol 57: 19-53.
- Kaufmann, H., X. Mazur, et al. (1999). "Influence of low temperature on productivity, proteome and protein phosphorylation of CHO cells." Biotechnol Bioeng 63(5): 573-82.
- Kim, N. S. and G. M. Lee (2000). "Overexpression of bcl-2 inhibits sodium butyrate-induced apoptosis in Chinese hamster ovary cells resulting in enhanced humanized antibody production." Biotechnol Bioeng 71(3): 184-93.
- Kim, V. N. and J. W. Nam (2006). "Genomics of microRNA." Trends Genet 22(3): 165-73.
- Kroemer, G. (2003). "Mitochondrial control of apoptosis: an introduction." Biochem Biophys Res Commun 304(3): 433-5.
- Krutzfeldt, J. and M. Stoffel (2006). "MicroRNAs: a new class of regulatory genes affecting metabolism." Cell Metab 4(1): 9-12.
- Kuchenbauer, F., R. D. Morin, et al. (2008). "In-depth characterization of the microRNA transcriptome in a leukemia progression model." Genome Res 18(11): 1787-97.
- Kumar, N., P. Gammell, et al. (2008). "Differential protein expression following low temperature culture of suspension CHO-K1 cells." BMC Biotechnol 8: 42.
- Kwaks, T. H., P. Barnett, et al. (2003). "Identification of anti-repressor elements that confer high and stable protein production in mammalian cells." Nat Biotechnol 21(5): 553-8.
- Lankat-Buttgereit, B. and R. Goke (2003). "Programmed cell death protein 4 (pdc4): a novel target for antineoplastic therapy?" Biol Cell 95(8): 515-9.
- Lee, R. C., R. L. Feinbaum, et al. (1993). "The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*." Cell 75(5): 843-54.
- Lewis, B. P., C. B. Burge, et al. (2005). "Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets." Cell 120(1): 15-20.
- Li, R., Y. Li, et al. (2008). "SOAP: short oligonucleotide alignment program." Bioinformatics 24(5): 713-4.
- Lim, L. P., N. C. Lau, et al. (2005). "Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs." Nature 433(7027): 769-73.

Lu, Z., M. Liu, et al. (2008). "MicroRNA-21 promotes cell transformation by targeting the programmed cell death 4 gene." Oncogene 27(31): 4373-9.

Mohan, C., Y. G. Kim, et al. (2008). "Assessment of cell engineering strategies for improved therapeutic protein production in CHO cells." Biotechnol J 3(5): 624-30.

Morin, R. D., M. D. O'Connor, et al. (2008). "Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells." Genome Res 18(4): 610-21.

Morlando, M., M. Ballarino, et al. (2008). "Primary microRNA transcripts are processed co-transcriptionally." Nat Struct Mol Biol.

Morozova, O. and M. A. Marra (2008). "Applications of next-generation sequencing technologies in functional genomics." Genomics 92(5): 255-64.

Muller, D., H. Katinger, et al. (2008). "MicroRNAs as targets for engineering of CHO cell factories." Trends Biotechnol 26(7): 359-65.

Mutskov, V. and G. Felsenfeld (2004). "Silencing of transgene transcription precedes methylation of promoter DNA and histone H3 lysine 9." EMBO J 23(1): 138-49.

Mutskov, V. J., C. M. Farrell, et al. (2002). "The barrier function of an insulator couples high histone acetylation levels with specific protection of promoter DNA from methylation." Genes Dev 16(12): 1540-54.

Nivitchanyong, T., A. Martinez, et al. (2007). "Anti-apoptotic genes Aven and E1B-19K enhance performance of BHK cells engineered to express recombinant factor VIII in batch and low perfusion cell culture." Biotechnol Bioeng 98(4): 825-41.

Orom, U. A., F. C. Nielsen, et al. (2008). "MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation." Mol Cell 30(4): 460-71.

Palermo, D. P., M. E. DeGraaf, et al. (1991). "Production of analytical quantities of recombinant proteins in Chinese hamster ovary cells using sodium butyrate to elevate gene expression." J Biotechnol 19(1): 35-47.

Pallavicini, M. G., P. S. DeTeresa, et al. (1990). "Effects of methotrexate on transfected DNA stability in mammalian cells." Mol Cell Biol 10(1): 401-4.

Peters, L. and G. Meister (2007). "Argonaute proteins: mediators of RNA silencing." Mol Cell 26(5): 611-23.

Puck, T. T. and F. T. Kao (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): 1227-34.

Richards, E. J. and S. C. Elgin (2002). "Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects." Cell 108(4): 489-500.

Ruby, J. G., C. H. Jan, et al. (2007). "Intronic microRNA precursors that bypass Drosha processing." Nature 448(7149): 83-6.

Sampson, V. B., N. H. Rong, et al. (2007). "MicroRNA let-7a down-regulates MYC and reverts MYC-induced growth in Burkitt lymphoma cells." Cancer Res 67(20): 9762-70.

Sandberg, R., J. R. Neilson, et al. (2008). "Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites." Science 320(5883): 1643-7.

Scorrano, L. and S. J. Korsmeyer (2003). "Mechanisms of cytochrome c release by proapoptotic BCL-2 family members." Biochem Biophys Res Commun 304(3): 437-44.

Sethupathy, P., B. Corda, et al. (2006). "TarBase: A comprehensive database of experimentally supported animal microRNA targets." RNA 12(2): 192-7.

Shibahara, K., M. Asano, et al. (1995). "Isolation of a novel mouse gene MA-3 that is induced upon programmed cell death." Gene 166(2): 297-301.

Si, M. L., S. Zhu, et al. (2007). "miR-21-mediated tumor growth." Oncogene 26(19): 2799-803.

Sunstrom, N. A., R. D. Gay, et al. (2000). "Insulin-like growth factor-I and transferrin mediate growth and survival of Chinese hamster ovary cells." Biotechnol Prog 16(5): 698-702.

Suzuki, C., R. G. Garces, et al. (2008). "PDCD4 inhibits translation initiation by binding to eIF4A using both its MA3 domains." Proc Natl Acad Sci U S A 105(9): 3274-9.

Taylor, M. W., J. H. Pipkorn, et al. (1977). "Purine mutants of mammalian cell lines: III. Control of purine biosynthesis in adenine phosphoribosyl transferase mutants of CHO cells." Somatic Cell Genet 3(2): 195-206.

Tey, B. T., R. P. Singh, et al. (2000). "Bcl-2 mediated suppression of apoptosis in myeloma NS0 cultures." J Biotechnol 79(2): 147-59.

Tey, B. T., R. P. Singh, et al. (2000). "Influence of bcl-2 on cell death during the cultivation of a Chinese hamster ovary cell line expressing a chimeric antibody." Biotechnol Bioeng 68(1): 31-43.

Trummer, E., K. Fauland, et al. (2006). "Process parameter shifting: Part I. Effect of DOT, pH, and temperature on the performance of Epo-Fc expressing CHO cells cultivated in controlled batch bioreactors." Biotechnol Bioeng 94(6): 1033-44.

Trummer, E., K. Fauland, et al. (2006). "Process parameter shifting: Part II. Biphasic cultivation-A tool for enhancing the volumetric productivity of batch processes using Epo-Fc expressing CHO cells." Biotechnol Bioeng 94(6): 1045-52.

- Urlaub, G. and L. A. Chasin (1980). "Isolation of Chinese hamster cell mutants deficient in dihydrofolate reductase activity." Proc Natl Acad Sci U S A 77(7): 4216-20.
- Wightman, B., I. Ha, et al. (1993). "Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*." Cell 75(5): 855-62.
- Wilczynska, A., N. Minshall, et al. (2009). "Two Piwi proteins, Xiwi and Xili, are expressed in the *Xenopus* female germline." RNA 15(2): 337-45.
- Wilson, T. J. and I. Kola (2001). "The LoxP/CRE system and genome modification." Methods Mol Biol 158: 83-94.
- Wurm, F. M. (2004). "Production of recombinant protein therapeutics in cultivated mammalian cells." Nat Biotechnol 22(11): 1393-8.
- Yallop, C. A., P. L. Norby, et al. (2003). "Characterisation of G418-induced metabolic load in recombinant CHO and BHK cells: effect on the activity and expression of central metabolic enzymes." Cytotechnology 42(2): 87-99.
- Yallop, C. A. and I. Svendsen (2001). "The effects of G418 on the growth and metabolism of recombinant mammalian cell lines." Cytotechnology 35(2): 101-14.
- Yang, H. S., A. P. Jansen, et al. (2001). "A novel transformation suppressor, Pcd4, inhibits AP-1 transactivation but not NF-kappaB or ODC transactivation." Oncogene 20(6): 669-76.
- Yang, H. S., J. L. Knies, et al. (2003). "Pcd4 suppresses tumor phenotype in JB6 cells by inhibiting AP-1 transactivation." Oncogene 22(24): 3712-20.
- Yang, H. S., C. P. Matthews, et al. (2006). "Tumorigenesis suppressor Pcd4 down-regulates mitogen-activated protein kinase kinase kinase 1 expression to suppress colon carcinoma cell invasion." Mol Cell Biol 26(4): 1297-306.
- Yoon, S. K., Y. H. Ahn, et al. (2007). "Effect of culture temperature on follicle-stimulating hormone production by Chinese hamster ovary cells in a perfusion bioreactor." Appl Microbiol Biotechnol 76(1): 83-9.
- Zanghi, J. A., W. A. Renner, et al. (2000). "The growth factor inhibitor suramin reduces apoptosis and cell aggregation in protein-free CHO cell batch cultures." Biotechnol Prog 16(3): 319-25.
- Zhu, S., M. L. Si, et al. (2007). "MicroRNA-21 targets the tumor suppressor gene tropomyosin 1 (TPM1)." J Biol Chem 282(19): 14328-36.
- Zhu, S., H. Wu, et al. (2008). "MicroRNA-21 targets tumor suppressor genes in invasion and metastasis." Cell Res 18(3): 350-9.

## 7. Appendix

### 7.1. Supplemental Table 1: Recombinant Protein Therapeutics produced in CHO

#	Product	Type	Therapeutic use	Manufacturer	Year of FDA approval
1	Activase	Tissue plasminogen activator	Acute myocardial infraction	Genentech	1987
2	Epogen/Procrit	Erythropoietin	Anemia	Amgen/OrthoBiotech	1989
3	Pulmozyme	Deoxyribonuclease I	Cystic fibrosis	Genentech	1993
4	Cerezyme	$\beta$ -glucocerebrosidase	Gaucher's disease	Genzyme	1994
5	Avonex	Interferon- $\beta$	Relapsing multiple sclerosis	BiogenIdec	1996
6	Follistim/Gonal-F	Follicle stimulating hormone	Infertility	Serono/NVOrganon	1997
7	Benefix	Factor IX	Hemophilia B	Wyeth	1997
8	Rituxan	Anti-CD20 mAb	Non-Hodgkin's lymphoma	Genentech, BiogenIdec	1997
9	Enbrel	TNF $\alpha$ receptor fusion	Rheumatoid arthritis	Amgen, Wyeth	1998
10	Herceptin	Anti-HER2 mAb	Metastatic breast cancer	Genentech	1998
11	Tenecteplase	Tissue plasminogen activator (engineered)	Myocardial infraction	Genentech	2000
12	ReFacto	Factor VIII	Hemophilia A	Wyeth	2000
13	Campath	Anti-CD52 mAb	Chronic lymphocytic leukemia	Genzyme, Bayer	2001
14	Aranesp	Erythropoietin (engineered)	Anemia	Amgen	2001
15	Humira	Anti-TNF $\alpha$ mAb	Rheumatoid arthritis	Abbott	2002
16	Rebif	Interferon- $\beta$	Relapsing multiple sclerosis	Serono	2002
17	Xolair	Anti-IgE mAb	Moderate/severe asthma	Genentech	2003
18	Fabrazyme	$\alpha$ -galactosidase	Fabry disease	Genzyme	2003
19	Advate	Factor VIII (engineered)	Hemophilia A	Baxter	2003
20	Raptiva	Anti-CD11a mAb	Chronic psoriasis	Genentech	2003
21	Avastin	Anti-VEGF mAb	Metastatic colorectal cancer & lung cancer	Genentech	2004
22	Luveris	Luteinizing hormone	Infertility	Serono	2004
23	Naglazyme	N-acetylgalactosamine-4-sulfatase	Mucopolysaccharidosis VI	BioMarin, Pharmaceutical	2005
24	Orencia	Ig-CTLA4 fusion	Rheumatoid arthritis	Bristol-Myers Squibb	2005
25	Vectibix	Anti-EGFR mAb	Metastatic colorectal cancer	Amgen	2006
26	Aldurazyme	Laronidase	Mucopolysaccharidosis I	Genzyme	2006
27	Myozyme	$\alpha$ -glucosidase	Pompe disease	Genzyme	2006