MicroRNA Profiling and Engineering of CHO Cell Lines Stably Expressing Difficult-to-Express Lysosomal Protein

Ifeanyi Amadi
MICRORNA PROFILING AND ENGINEERING OF CHO CELL LINES STABLY EXPRESSING DIFFICULT-TO-EXPRESS LYSOSOMAL PROTEIN

BY

IFEANYI MICHAEL AMADI

A Dissertation submitted to the Faculty of Keck Graduate Institute of Applied Life Sciences in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Applied Life Sciences

Claremont, California
2018

Approved by:

Parviz Shamlou, Chair

Copyright by Ifeanyi Michael Amadi. 2018
All Rights Reserved
We, the undersigned, certify that we have read this dissertation of Ifeanyi Michael Amadi and approve it as adequate in scope and quality for the degree of Doctor of Philosophy.

Dissertation Committee:

Parviz Shamlou, Chair

Jon LeBowitz, Member

Terri Christianson, Member

Vishal Agrawal, Member

Cameron Bardliving, Member, PhD Program Director
MicroRNA Profiling and Engineering of CHO Cell Lines Stably Expressing Difficult-To-Express Lysosomal Protein

By:

IFEANYI MICHAEL AMADI

Keck Graduate Institute, 2018

ABSTRACT

Difficult-to-express (DTE) recombinant proteins like multi-specific proteins, DTE monoclonal antibodies and lysosomal enzymes, have seen difficulties in manufacturability using Chinese hamster ovary (CHO) cells and other mammalian cells as production platforms. CHO cells are preferably used for protein production because of their innate ability to secrete human-like recombinant proteins with post-translational modification, resistance to viral infection and familiarity with drug regulators. However, despite huge progress made in engineering CHO cells for high volumetric productivity, expression of DTE proteins like recombinant lysosomal sulfatase represent one of the poorly understood proteins. Furthermore, there are growing interest in the use of microRNAs (miRNAs) to engineer CHO cells expressing DTE proteins to improve cell performance of relevant bioprocess phenotypes. Therefore, we sought to understand miRNA expression profiles in CHO cell lines stably expressing DTE lysosomal protein and examined the effect of microRNA-engineering of these cell lines on protein expression.

Firstly, we utilized next generation sequencing (NGS) technology for an integrated mRNA and microRNA profiling of three CHO cell lines (including parental cell) stably expressing a DTE lysosomal protein cultivated in a biphasic fed-batch mode within a 5L Dasgip bioreactor. With
exception to the parental cell line, the other two cell lines differ by their productivity and were identified as low and high producers. Following RNA-seq and small RNA-seq data analyses by the Bioinformatics group at BioMarin, among other analyses, correlation and differential analyses identified 500 mRNA and 35 miRNAs that were differentially expressed in these cell lines over cultivation period. This study identified miRNAs that are potential targets for engineering of CHO cells for the enhancement of DTE protein expression. Secondly, in an independent study, we identified miR-23a and miR-377 as miRNAs targeting sulfatase modifying factor 1 (SUMF1) by using in silico prediction tools as SUMF1 is an activator of sulfatases. Transient inhibition of endogenous miR-23a/miR-377 significantly enhanced recombinant sulfatase enzyme specific activity in CHO cells without affecting cell growth. Though, inhibition of miR-23a/miR-377 had no significant effect on the mRNA and protein levels of SUMF1, overexpression of miR-23a/377 significantly reduced both the mRNA and protein levels of SUMF1. In summary, our data demonstrates the importance of using miRNA to optimize protein expression in CHO cells secreting DTE recombinant lysosomal proteins.
DEDICATION

To my family and friends, most especially to my mother, Lovender Amadi, who made me who I’m today and to my wonderful wife, Amaka, for her unconditional love and dedication and being there for me.
ACKNOWLEDGMENTS

I want to use this opportunity to express my sincere gratitude to my supervisors (Terri Christianson, Vishal Agrawal and Jon LeBowitz) at BioMarin Pharmaceutical for granting me the opportunity to pursue this research and providing time from their busy schedules to attend meetings on this research. I can’t thank you all enough for your confidence in me and continued advice throughout the process. I wish words could explain my feelings and genuine gratitude as I write this piece.

Also, I would like to say a big thank you to my professors at the Keck Graduate Institute (KGI), Parviz Shamlou and Cameron Bardliving, for their advice and feedbacks during this period. In addition, thank you for granting me the opportunity to present my work at several occasions at KGI.

In addition, I want extend a much gratitude to my fellow lab mates at BioMarin for their continued encouragement and collaborative work style.

Lastly, I would like to send out my loving appreciation to my wife, Amaka, for her obvious love and understanding throughout this period and how she tells me that the only word she kept on seeing on my books any time she steps close to my reading table is “microRNA”.

Contents

CHAPTER 1 .......................................................................................................................................................... 1

1.0 INTRODUCTION AND LITERATURE REVIEW ...................................................................................... 1

1.1 Bioprocessing - production of therapeutic proteins ............................................................................. 3

1.2 MicroRNA .................................................................................................................................................. 3

1.2.1 MiRNA Biogenesis: .......................................................................................................................... 5

1.2.2 Mechanism of action ......................................................................................................................... 6

1.2.3 Methods of engineering mammalian cells using miRNA ............................................................. 7

1.2.3.1 mRNA and miRNA profiling of CHO cells .............................................................................. 7

1.2.4 MicroRNA-engineering of CHO cells ............................................................................................. 10

1.2.4.1 MicroRNA engineering of CHO cells expressing difficult-to-express (DTE) protein .................. 10

1.3 Difficult-to-express lysosomal enzymes .............................................................................................. 11

CHAPTER 2 ................................................................................................................................................... 14

2.0 RESEARCH OBJECTIVES ....................................................................................................................... 14

CHAPTER 3 ..................................................................................................................................................... 15

3.0 MATERIALS AND METHODS .............................................................................................................. 15

3.1 CELL CULTURE TECHNIQUES AND TRANSFECTION ...................................................................... 15

3.1.1 Biosafety cabinet ............................................................................................................................... 15

3.1.2 Cell Thawing .................................................................................................................................... 15
3.1.3 Cell Freezing ................................................................. 15

3.2 Passaging of cell lines .................................................... 16

3.2.1 Adherent cells ......................................................... 16

3.2.2 Suspension cells ...................................................... 16

3.2.3 Cell counting .......................................................... 17

3.2.3.1 Vicell™ XR cell viability analyzer ............................... 17

3.2.3.2 Nucleocounter NC-200 ........................................... 17

3.3 Transfection ............................................................. 17

3.3.1 Transfection of siRNA/microRNAs with lipid-base reagent method .......................... 17

3.3.2 Transfection of plasmid DNA (pDNA) using electroporation method .... 18

3.3.3 Cell Imaging/Transfection efficiency measurement ........................ 19

3.4 MOLECULAR BIOLOGY TECHNIQUES ................................. 20

3.4.1 Molecular biology water .............................................. 20

3.4.2 Sterilization ............................................................ 20

3.4.3 Reconstitution of lyophilized miRNA ................................ 20

3.4.4 Rehydration of antibodies ........................................... 20

3.4.5 Extraction of total RNA (including small RNA - microRNA) .............. 21

3.4.6 Determination of nucleic acid concentration ........................ 22

3.4.6.1 Plasmid DNA ...................................................... 22

3.4.6.2 Total RNA ........................................................ 22
3.4.7 Determination of RNA quantity and quality using Bioanalyzer ............ 22
3.4.8 Reverse transcription (cDNA generation) ..................................... 24
3.4.9 Real Time-PCR (qPCR) ............................................................... 24
3.4.10 Molecular cloning ................................................................. 26
3.4.10.1 Agarose gel preparation ....................................................... 26
3.4.10.2 Gel extraction ................................................................. 26
3.4.10.3 PCR purification .............................................................. 27
3.4.10.4 Plasmid DNA purification (maxiprep) .................................. 28
3.4.10.5 Transformation ................................................................. 28
3.4.10.6 Purification of plasmid DNA (miniprep) ................................ 29
2.4.10.7 Restriction endonuclease digest ......................................... 29
3.4.10.8 BigDye® Polymerase Chain Reaction (PCR) cycle sequencing .... 30
3.4.10.9 DNA sequencing clean-up ............................................... 31
3.5 PROTEIN ANALYSIS ............................................................... 32
3.5.1 Bicinchoninic acid (BCA) Protein Assay (microplate procedure) ...... 32
3.5.2 Western Blot (Immunoblotting) ............................................... 32
CHAPTER 4 .................................................................................. 34
4.0 Introduction ........................................................................... 37
4.1 Materials and Methods .......................................................... 39
4.1.1 Cell lines and cell culture ...................................................... 39
4.1.2 Bioreactor fed-batch cultivation .......................................................... 40
4.1.3 Transient transfection of mimics and inhibitors ....................................... 41
4.1.4 Real-time Quantitative Reverse-Transcription PCR (qRT-PCR) ............. 41
4.1.5 Western blot analysis .................................................................................. 42
4.1.6 Enzymatic assay (titer) ................................................................................. 42
4.1.7 Bioinformatics analyses .............................................................................. 43
4.2 Results ............................................................................................................. 43
4.2.1 Fed-batch cell culture cultivation ............................................................... 43
4.2.2 Identification of growth-correlating miRNAs and mRNAs ....................... 44
4.2.3 Identification of titer-correlating miRNAs and mRNAs ............................. 45
4.2.4 Functional confirmation of miRNA expression by qRT-PCR ................. 45
4.2.5 Functional effect of miR-31-5p on the secreted lysosomal protein of 
CHO-lys-low and CHO-lys-high producers .................................................... 46
4.2.6 Identification of miR-31-5p target genes and functional annotation ...... 46
4.3 Discussion ........................................................................................................ 47

CHAPTER 5 ............................................................................................................. 78
5.0 Introduction ..................................................................................................... 80
5.1 Materials and Methods .................................................................................. 82
5.1.1 Cell culture ................................................................................................ 82
5.1.2 Transfection ......................................................... 83
5.1.3 RNA Isolation ............................................................ 83
5.1.4 Real-time Quantitative Reverse-Transcription PCR (qRT-PCR) ............ 84
5.1.5 Western blot analysis ...................................................... 84
5.1.6 Sulfatase enzymatic activity and Elisa assays .................................. 85
5.1.7 Bioinformatics .................................................................. 86
5.2 Results .............................................................................. 86
5.2.1 In silico identification of putative miRNAs targeting SUMF1 ................ 86
5.2.2 Characterization of different mammalian cells to determine appropriate SUMF1 antibody ................................. 86
5.2.3 Transient transfection of miR-23a and miR-377 mimics reduced SUMF1 protein and mRNA levels with no effect on sulfatase activity .................... 87
5.2.4 Transient inhibition of miR-23a and miR-377 enhances recombinant sulfatase enzyme specific activity in CHO cell line ........................................... 88
5.3 Discussion .......................................................................... 89
5.3.1 Transient inhibition of miR-23a and miR-377 increases lysosomal recombinant sulfatase specific activity in CHO cell ............................................ 91
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>LC/MS</td>
<td>Liquid chromatography/mass spectrometry</td>
</tr>
<tr>
<td>ER</td>
<td>Endoplasmic reticulum</td>
</tr>
<tr>
<td>miRNA</td>
<td>microRNA</td>
</tr>
<tr>
<td>miR</td>
<td>matured miRNA</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>3' UTR</td>
<td>3' Untranslated region of mRNA</td>
</tr>
<tr>
<td>Pri-miRNA</td>
<td>Primary microRNA</td>
</tr>
<tr>
<td>Pre-miRNA</td>
<td>Precursor microRNA stem loop</td>
</tr>
<tr>
<td>DGCR8</td>
<td>DiGeorge critical syndrome 8</td>
</tr>
<tr>
<td>miRISC</td>
<td>miRNA-induced silencing complex</td>
</tr>
<tr>
<td>FGE</td>
<td>Formyl-glycine generating (FGly) enzyme</td>
</tr>
<tr>
<td>SUMF1</td>
<td>Sulfatase modifying factor 1 (SUMF1)</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td>Real-time quantitative reverse transcriptase PCR</td>
</tr>
<tr>
<td>AntimiRs</td>
<td>miRNA inhibitors</td>
</tr>
<tr>
<td>DTE</td>
<td>Difficult-to-express</td>
</tr>
<tr>
<td>PTM</td>
<td>Post-translational modification</td>
</tr>
<tr>
<td>Si-RNA</td>
<td>Small interfering RNA</td>
</tr>
<tr>
<td>NT-SiRNA</td>
<td>Non-targeting siRNA</td>
</tr>
<tr>
<td>CHO</td>
<td>Chinese hamster ovary</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>HEK</td>
<td>Human embryonic kidney</td>
</tr>
<tr>
<td>MCF7</td>
<td>Human breast cancer cell</td>
</tr>
<tr>
<td>DE</td>
<td>Differentially expressed</td>
</tr>
<tr>
<td>CHO-lys-low</td>
<td>Low producer CHO cell secreting a lysosomal enzyme</td>
</tr>
<tr>
<td>CHO-lys-high</td>
<td>High producer CHO cell secreting a lysosomal enzyme</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RNA-Seq</td>
<td>RNA sequencing</td>
</tr>
<tr>
<td>DAVID</td>
<td>Database for Annotation, Visualization and Integrated Discovery</td>
</tr>
<tr>
<td>mmu</td>
<td>Mus musculus (mouse)</td>
</tr>
<tr>
<td>FDR</td>
<td>False discovery rate</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

**Figure 1**: Canonical pathway of microRNA biogenesis. Image used and adapted from Winter et al. (2009)(Winter et al. 2009) .............................................................................................................................................6

**Figure 2**: Growth profiles of CHO-lys-low producer, CHO-lys-high producer and CHO-non-producer (parental cell line) grown in 5L bioreactors in 3L w/v cell culture for 14 days (A) specific growth rate (B) Viable cell concentration (C). Viability (D) Integral viable cell density (IVCD) ........................................69

**Figure 3**: Productivity profiles and offline Glucose concentration/pH trends of CHO-lys-low producer, CHO-lys-high producer and CHO-non-producer (parental cell line) grown in 5L bioreactors in 3L w/v cell culture for 14 days (A) Specific productivity (B) Titer (C) Glucose concentration (D) Offline pH ........................................................................................................................................70

**Figure 4**: qRT-PCR confirmation of 14 DE (FDR<0.05 & FC>2 or <0.5) miRNAs selected from 35 DE (FDR<0.5) miRNAs identified by small RNA-Seq over cell culture duration. MiRNA expression was normalized to U6 snoRNA. Error bars represent the standard error of mean (SEM) of three independent experiments (n=3). For statistical analysis, one-way ANOVA comparing the mean values of the mimics (* P < 0.05; ** P < 0.01; *** P < 0.001). ........................................................................................................72

**Figure 5**: Quantitative real-time PCR (qRT-PCR) analyses. Cells were transiently transfected with 50 nM of miR-31-5p mimics or inhibitors in CHO-lys-low producer expressing DTE lysosomal enzyme. Cell lysates harvested 48 h post-transfection (A) Endogenous miR-31-5p levels post-transfection with miR-31-5p inhibitor (B) Endogenous miR-31-5p levels post-transfection of miR-31-5p mimics .........................................................................................................................73

**Figure 6**: Growth profile of CHO-lys-low producer post-transfection with (A) mouse miR-31-5p mimic (B) mouse miR-31-5p inhibitor (C) Product concentration (titer) of lysosomal enzyme secreted by CHO-lys-low producer post-transfection of mouse miR-31-5p inhibitor (D) Product concentration (titer) of lysosomal enzyme secreted by CHO-lys-low producer post-transfection of mouse miR-31-5p mimic (E) Product concentration (titer) of lysosomal enzyme secreted by CHO-lys-high producer post-transfection of mouse miR-31-5p mimic and inhibitor. Error bars represent the standard error of mean (SEM) of three independent experiments (n=3). Error bars represent the standard error of mean (SEM) of three independent experiments (n=3). For statistical analysis, student t-test was used (* P < 0.05; ** P < 0.01; *** P < 0.001) .........................................................................................................................74

**Figure 7**: (A) Functional annotation clustering of DE target genes of miR-31-5p predicted by at least 3 in silico tools in combination with miRNA-target gene result from the NGS data analyses. (B)
Quantitative real-time PCR (qRT-PCR) analyses of day 3 RNA samples used for NGS for both CHO-lys-low and CHO-lys-high producer cell lines. mRNA levels of GALC.

**Figure 8:** Characterization of different mammalian cells to determine appropriate SUMF1 antibody (A) Western blot detection of endogenous SUMF1 protein 72 h post transfection. In addition, SUMF1 protein detection in CHO-sulfatase cell stably expressing human SUMF1 was also tested (B) Predicted binding sites of miR-23a and miR-377 on SUMF1 3’ UTR. Bases in red represents the binding site of the miRNAs seed region to the 3’ UTR of SUMF1 (TargetScan.org).

**Figure 9:** Transient transfection of 50 nM miR-23a and miR-377 mimics in CHO suspension cells stably expressing a recombinant difficult-to-express (DTE) lysosomal sulfatase protein. Cells were cultured in a 6-well plate using a batch process and kept in a static incubator. (A) Viable cell density (VCD) and viability 72 h post transfection (B) Sulfatase titer and (C) specific activity determined by an Elisa and activity assay respectively. Samples were taken from the culture supernatants 72 h post-transfection and data represented as fold change normalized to scramble (NT-siRNA) control. Anti-SUMF1 siRNA (si-SUMF1) was used as a positive control. Error bars represent the standard error of mean (SEM) of three independent experiments (n=3). For statistical analysis, one-way ANOVA comparing the mean values of the mimics (** P < 0.001).

**Figure 10:** Functional validation of SUMF1 as a putative target of miR-23a and miR-377 using Western blot and quantitative real-time PCR (qRT-PCR) analyses. Cells were transiently transfected with 50 nM of miR-23a and miR-377 mimics in CHO suspension cells stably expressing a recombinant difficult-to-express (DTE) lysosomal sulfatase protein and cell lysates harvested 48 h post-transfection. (A) Western blot showing SUMF1 protein levels and relative densitometry signal values of SUMF1 protein levels relative to GAPDH (B) Endogenous SUMF1 mRNA levels post-transfection relative to scramble (NT-siRNA) non-targeting control. SUMF1 expression was normalized to GAPDH (C) Endogenous miR-23a and miR-377 levels post-transfection relative to scramble. MiRNA expression was normalized to U6 snoRNA. Anti-SUMF1 siRNA (si-SUMF1) was used as a positive control. Error bars represent the standard error of mean (SEM) of three independent experiments (n=3). For statistical analysis, one-way ANOVA comparing the mean values of the mimics (* P < 0.05; ** P < 0.01; *** P < 0.001).

**Figure 11:** Transient transfection of 50 nM miR-23a and miR-377 antimiRs in (targeting either endogenous mouse or CHO miR-23a or miR-377) CHO suspension cells stably expressing a recombinant difficult-to-express (DTE) lysosomal sulfatase protein. Cells were cultured in a 6-well plate using a batch process and kept in a static incubator. (A) Viable cell density (VCD) and viability 72 h post transfection (B) Sulfatase titer and (C) specific activity determined by an Elisa and activity assay respectively. Samples were taken from the culture supernatants 72 h post-transfection and
data represented as fold change normalized to scramble (NT-siRNA) control. Anti-SUMF1 siRNA (si-SUMF1) was used as a positive control. Error bars represent the standard error of mean (SEM) of three independent experiments (n=3). For statistical analysis, one-way ANOVA comparing the mean values of the (* P < 0.05; ** P < 0.01; *** P < 0.001).

**Figure 12:** Effect of transient transfection of 50 nM miR-23a and miR-377 antimiRs in (targeting either endogenous mouse or CHO miR-23a or miR-377) CHO suspension cells stably expressing a recombinant difficult-to-express (DTE) lysosomal sulfatase protein using Western blot and quantitative real-time PCR (qRT-PCR) analyses. (A) Western blot of SUMF1 protein levels and relative densitometry signal values of SUMF1 protein levels relative to GAPDH 48 h post transfection (B) Endogenous SUMF1 mRNA levels post-transfection relative to scramble (NT-siRNA) non-targeting control. SUMF1 expression was normalized to GAPDH (C) Endogenous miR-23a and miR-377 levels 48 h post-transfection relative to scramble. miRNA expression was normalized to U6 snoRNA. Anti-SUMF1 siRNA (si-SUMF1) was used as a positive control. Error bars represent the standard error of mean (SEM) of three independent experiments (n=3). Error bars represent the standard error of mean (SEM) of three independent experiments (n=3). For statistical analysis, one-way ANOVA comparing the mean values of the (* P < 0.05; ** P < 0.01; *** P < 0.001).

**Figure 13:** Transient transfection of titrated concentration (10-200 nM) of cgr-miR-23a mimics in CHO suspension cells stably expressing a recombinant difficult-to-express (DTE) lysosomal sulfatase protein. Cells were cultured in a 6-well plate using a batch process and kept in a static incubator. (A) Upper panel: Viable cell density (VCD). Lower Panel: Viability. Both 72 h post-transfection (B) Western blot of SUMF1 protein levels and densitometry signal values of SUMF1 protein levels relative to GAPDH 48 h post transfection (C) Endogenous SUMF1 mRNA levels post-transfection relative to scramble (NT-siRNA) non-targeting control. SUMF1 expression was normalized to GAPDH (D) Endogenous miR-23a levels 48 h post-transfection relative to scramble. miRNA expression was normalized to U6 snoRNA. Anti-SUMF1 siRNA (si-SUMF1) was used as a positive control. Data are mean +/- SEM of two independent experiment. For statistical analysis, unpaired two-tailed t-test was applied (* P < 0.05; ** P < 0.01).

**Figure 14:** si-hSUMF1 was titrated to determine the optimal concentration that will reduce or completely delete exogenously hSUMF1 stably expressed in CHO cell line secreting a sulfatase: A) Western blot analysis of different concentration of si-hSUMF1 and volume of transfection reagent used. B) Relative signal intensity of hSUMF1 levels in CHO cell line exogenously expressing hSUMF1 compared to mock, scramble and untransfected cells.
LIST OF TABLES

Table 1: Advantages and disadvantages of miRNA profiling technologies and their commercially available platforms* ....................................................................................................................... 8

Table 2: mRNA and miRNA profiling experiment conducted on CHO cell host platforms .......... 9

Table 3: Reaction master mix for the detection of matured miRNA in a 384-well plate ............. 25

Table 4: Roche LightCycler 480 cycling conditions ............................................................................. 25

Table 5: BigDye reaction mix preparation ............................................................................................. 31

Table 6: BigDye PCR reaction parameters ............................................................................................. 31

Table 7: Parameters used for next-generation sequencing................................................................. 64

Table 8: Relevant bioprocess characteristics of CHO cell lines used for cell culture in 5L bioreactor .............................................................................................................................................. 64

Table 9: Growth correlating DE miRNAs unique to CHO-lys-high producer ................................. 65

Table 10: Titer correlating DE miRNAs unique to CHO-lys-high producer ................................. 65

Table 11: Gene ontology and Functional annotation of miR-31-5p target genes predicted by at least 3 in silico tools in combination with DE genes from the RNA-Seq data along with potential genes for downstream analyses ........................................................................................................................................ 65

Table 12: Genes selected for downstream analysis from the functional annotation result of miR-31 target genes derived from prediction by at least 3 in silico tools and DE genes from RNA-Seq data. ........................................................................................................................................ 66

Table 13: Functions of potential miR-31-5p target genes with relevance to bioprocess phenotypes ........................................................................................................................................ 67

Table 14: Primers used for miRNA and mRNA expression analyses using qRT-PCR ............... 76

Table 15: miRTarbase (v 7.0) list of validated targets of positively titer-correlating miRNAs unique to CHO-lys-high producer cell line ........................................................................................................ 76

Table 16: Primers used for miRNA and mRNA expression analyses using qRT-PCR ............. 108

Table 17: MiRNAs predicted to target SUMF1 gene in mouse ......................................................... 109
CHAPTER 1

1.0 INTRODUCTION AND LITERATURE REVIEW

The increase in production of more complex, difficult-to-express (DTE) recombinant therapeutic proteins e.g. multi-specific proteins, fusion proteins, membrane proteins, lysosomal enzymes and other DTE monoclonal antibodies (mAb) have seen difficulties in manufacturability using CHO cells or other mammalian cells as production factories (Kelly et al. 2014; Migani, Smales, and Bracewell 2017). Mammalian cells are preferably used for the production of most proteins because of their ability to produce proteins with human-like post-translation modification (PTM), their resistance to viral infection and popularity with drug regulators. Lysosomal enzymes represent one of the least studied or reviewed recombinant therapeutic class of proteins compared to monoclonal antibodies (mAbs) and other well understood proteins manufactured as biologics. In the early years of lysosomal enzyme production as a biopharmaceutical, CHO cells production of lysosomal enzyme could be as low as 20,900 U/mg (~100 micrograms of enzyme/10^7 cells) produced, secreting ~ 13,000 U (or 75 µg/10^7 cells) per day (Ioannou, Bishop, and Desnick 1992). Of recent, the productivity yield has increased (e.g. 0.05-0.25 mg/mL of GAA production) as a result of advances in bioprocessing process parameter control, media formulation and genetic engineering (Migani, Smales, and Bracewell 2017). In spite of these advances the full potential of the mammalian host systems used in the production of DTE lysosomal proteins have not been completely utilized. At this time point, engineering of the mammalian host system might represent a viable route to increasing production yield of DTE lysosomal proteins. The lysosome which acts like the recycle warehouse of the cell, contains proteins that breakdown glycosaminoglycans (GAG) and other macromolecular substrates. Mutations in the gene coding for any of these
lysosomal enzymes or proteins can lead to the accumulation of these substrates in the lysosome leading to a disease state termed lysosomal storage disorder (LSD). An example of these diseases among others is multiple sulfatase deficiency (MSD) caused by mutations in sulfatase modifying factor (SUMF1), a gene that encodes formylglycine –generating enzyme (FGE). A difficult-to-express protein could arise as a result of insolubility, improper folding, aggregate formation (e.g. variable region of dimeric Fab antibody), and issues stemming from protein half-life, production yield and stability. In the world of rare diseases – lysosomal storage disorders (LSD), producing active form of lysosomal protein could sometimes pose a challenge to scientists.

The role of microRNA in cancer and autoimmune diseases is well researched unlike its application in bioprocessing for protein expression purposes. MicroRNAs (miRNAs) has been shown to play critical roles in many cellular processes – differentiation, development, cellular growth, apoptosis, including the expression of transcription factor EB (TFEB) that regulates a gene network called Coordinated Lysosomal Expression and Regulation (CLEAR) (de Queiroz et al. 2016). A handful of work has been done linking non-coding RNA to gene expression in lysosomal storage diseases (LSD). Generally, microRNAs can be located between exons or introns of the genome. They tend to destabilize or inhibit the translation of their mRNA target by imperfectly binding to the 3’ UTR of the mRNA target. MicroRNA effect on CHO cells that has been a subject of intense research currently and it has been observed that as CHO cells behavior defer by the kind of protein they express and/or clonal variation, miRNA expression profiles changes (Maccani et al. 2014). Diiferent methods have been developed for the identification of miRNA expression relative to protein expression: microRNA library screen, mRNA and miRNA microarray, next-generation sequencing (Inwood, Betenbaugh, and Shiloach 2018).
1.1 Bioprocessing - production of therapeutic proteins

Recombinant therapeutic proteins are usually produced in large scale via a process called bioprocessing. Bioprocessing, in the context of protein production, entails the use of living organisms to produce recombinant therapeutic proteins – antibodies, fusion proteins, anticoagulants, growth factors, hormones, interferons, enzymes, vaccines. The protein production process involves multiple stages from selection of expression host system to the use bioreactors for production. The choice of expression host like mammalian cells (e.g., Chinese hamster ovary cell (CHO), human embryonic kidney cells (HEK), Escherichia coli, insect cells, yeast, hybridomas etc.) depends on: the type of biopharmaceutical product (e.g. membrane protein, glycosylated proteins), post-translation modification (PTM) requirement, ease of use, protein stability and solubility, timing, cost, and familiarity with regulators among others.

1.2 MicroRNA

MicroRNAs (miRNAs) are short non-coding ribonucleic acid (RNA) molecule, about 18-24 nucleotide in length. They negatively regulate genes post-transcriptionally, regulating one-third of the genes in eukaryotic cells without adding any translational burden on the cell machineries unlike protein coding genes when overexpressed. MiRNAs regulate genes by binding imperfectly (one or two base pair difference) to the 3' untranslated region (UTR) of their target mRNA with mismatches and bulges causing repression or a perfect binding causing degradation by the miRNA-induced silencing complex (miRISC) (Ha and Kim 2014). The 5' end of miRNA at nucleotide 2 to 8 contain a domain region called “seed region” for the recognition of target mRNA. Though in plants, microRNAs bind to their targets in a near-
perfect manner inducing the cleavage of mRNA targets by endonuclease (Pasquinelli 2012). Though a rare example of near-perfect complementarity has been mentioned for animals in which miR-196 binds to HOXB8 mRNA (Pasquinelli 2012).

One major feature of MicroRNAs is their known conservative characteristics across species, and even as similar target recognition seem to be conserved in mammalian cells (Friedman et al. 2009). MiRNAs can either be downregulated or overexpressed where alterations in their levels can cause changes in cell phenotypes. They have been described as molecules that fine tune gene expression in eukaryotic cells (Baek et al. 2008) and have been shown to play vital roles in cellular processes like differentiation, proliferation, development, and apoptosis. Alterations in the levels of miRNA has been linked to various diseases, in most cases cancer (Wang, Wei, and Sarkar 2012). These findings are prominent in cancer research for over two decades now and protein expression scientists are currently exploring ways to tap into the potentials of miRNAs as global gene regulators in the field of bioprocessing. Techniques like crosslinking and immunoprecipitation (CLIP) and ribosome profiling have been employed to understand miRNA binding sites and its targets (Pritchard, Cheng, and Tewari 2012). Different culture conditions (nutrient depletion, temperature shift) has been shown to cause changes in miRNAs levels in profiling studies in a bioprocessing system (Druz et al. 2013; N. Barron et al. 2011).

For example, Strotbek et al (2013) were able to demonstrate that miR-557 and miR-1287 improved growth profile (VCD) and specific productivity (qP) respectively in CHO cells expressing IgG1 proteins. The CHO cells engineered with these miRNAs exhibited higher IgG1 titer while retaining product quality. In addition, the CHO cells exhibited a much higher protein titer and specific productivity when the miRNAs were co-transfected (Strotbek et al. 2013).
1.2.1 MiRNA Biogenesis:

MicroRNAs are transcribed by Polymerase II (Pol II) into a primary transcript (pri-miRNA) that is about 72 nucleotide in length inside the nucleus of the cell (Figure 1). Thereafter, it is processed by a microprocessor complex- Drosha and its cofactor DiGeorge syndrome critical region gene 8 (DGCR8 or Pasha), into a hairpin precursor microRNA (pre-miRNA) which is then transported into the cytoplasm by RanGTP-dependent exportin-5 (XPO5) protein. While in the cytoplasm, the loop of the pre-miRNA is removed by DICER/Tar RNA binding protein (TRBP) complex forming a microRNA duplex- guide and passenger (or star) strands that is 21-24 nucleotide long. The resulting microRNA duplex is later bound by argonaute protein forming the miRNA-induced silencing complex (miRISC), includes GW182 proteins, which is directed by the guide strand to mRNA target for regulation. The star strand (miRNA*) is degraded and the guide strand binds to the target mRNA through partial complementarity with bulges resulting in repression of translational initiation or cleavage of the target mRNA in some rare occasion of near perfect complementarity e.g., miR-196 targeting HOXB8 mRNA (Pasquinelli 2012). The selection of the strand destined for degradation is dependent on the thermodynamic stability of the 5’ end (Pasquinelli 2012; O’Brien et al. 2018; Ha and Kim 2014).
1.2.2 Mechanism of action

The effect of synthetic microRNA inhibitors on cell physiology is dependent on abundance of endogenous microRNA expression it is inhibiting. MicroRNA recognizes its target by binding to the 3’ UTR of the mRNA target. Though there are reports that revealed binding to the mRNA target at the 5’ UTR (J. Robin Lytle, Therese A. Yario 2000), protein coding region,
and 3' UTR. Though there are reports that miRNAs can also upregulate the expression of their target genes as against their nominal function of negatively inhibit gene expression (Vasudevan 2012).

1.2.3 Methods of engineering mammalian cells using miRNA

In most cases, transient transfection of microRNA mimics or inhibitors into mammalian cells is performed by lipid-mediated transfection method. To efficiently assess transfection efficiency of the transfection reagent used it is pertinent to assess the effect of the miRNA mimic or inhibitor on their target mRNA/protein in addition to using flow cytometry or fluorescence microscopy. The reason is that using flow cytometry or fluorescence microscopy alone could give a false positive result because the transfected microRNAs could be stalked in the endosome, firmly bound with the transfection reagent and not being released in order to be bound with the miRISC complex for mRNA targeting (Fischer et al. 2013). Interestingly, assessing transfection efficiency by measuring the levels of miRNA mimics using qRT-PCR might not really portray the accurate amount of functional or active miRNAs bound to Argonaute protein, while microRNA inhibitors could inhibit qPCR reaction. Using qRT-PCR alone might only be measuring the amount of both vesicular (inactive) and Argonaute-bound (active) mimics (Thomson et al. 2013). Thomson et al. 2013 used Argonaute:miRNA immunoprecipitation assay to accurately assess functional or active microRNAs post transient transfection of microRNA mimics.

1.2.3.1 mRNA and miRNA profiling of CHO cells

To generate a better understanding of the transcriptome (mRNA or miRNA) of CHO cells, scientists are using high throughput technologies (e.g. next-generation sequencing, microarray technology, liquid chromatography/mass spectrometry (LC/MS) and miRNA library) in combination with cell cultivation (suspension or static) and different cell culture
modes (fed-batch, batch or steady state) (Table 2). Only of recent was the CHO genome was sequenced providing more opportunities to understand the biology of this important host cell. Profiling of CHO cells using these technologies have providing much characterization of CHO cell lines and context for cell engineering purposes. Although, these profiling techniques have their advantages and disadvantages (Table 1) (Pritchard, Cheng, and Tewari 2012)

Table 1: Advantages and disadvantages of miRNA profiling technologies and their commercially available platforms*

<table>
<thead>
<tr>
<th>MiRNA Profiling Technologies</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>Platforms</th>
</tr>
</thead>
<tbody>
<tr>
<td>MicroRNA microarray</td>
<td>Low cost and high throughput</td>
<td>Lower specificity compared to qRT-PCR or RNA sequencing</td>
<td>Genome Biochip miRNA, GeneChip miRNA array, GenoExplorer, MicroRNA microarray, miRCURY LNA microRNA array, OneArray, uParaFlo biochip array</td>
</tr>
<tr>
<td></td>
<td>Affordable</td>
<td>Cross hybridization between probes and sample</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No detection of novel miRNAs</td>
<td></td>
</tr>
<tr>
<td>RNA sequencing</td>
<td>No cross hybridization between samples</td>
<td>Required expertise for computational analyses of data</td>
<td>HTS: HiSeq 2000, SOLID, GS FLX+ (454 sequencing)</td>
</tr>
<tr>
<td></td>
<td>Higher sensitivity detecting very less abundant genes (Stiefel et al. 2016)</td>
<td>Cannot be used for absolute quantification</td>
<td>Small-scale: Ion Torrent, MiSeq, GS Junior (454)</td>
</tr>
<tr>
<td></td>
<td>Broader dynamic range. Can detect novel miRNAs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>qRT-PCR</td>
<td>No cross hybridization between samples</td>
<td>Cannot identify novel miRNAs</td>
<td>Taqman individual assays, miRCURY LNA qPCR, TaqMan TLDA microfluidics card, Biomark HD system, SmartChip human microRNA, miScript miRNA PCR array</td>
</tr>
<tr>
<td></td>
<td>High sensitivity and specificity</td>
<td>Medium throughput with respect to number of samples processed per day</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Can be used for absolute quantification</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table modified from Pritchard, Cheng and Tewari (Pritchard, Cheng, and Tewari 2012)
### Table 2: mRNA and miRNA profiling experiment conducted on CHO cell host platforms

<table>
<thead>
<tr>
<th>Profiling Technique</th>
<th>Cell Line Used</th>
<th>Cell Type</th>
<th>Protein Secreted</th>
<th>Cultivation mode</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human microRNA library</td>
<td>CHO-DG44</td>
<td>Suspension</td>
<td>Human serum albumin (HSA) or mAb IgG1</td>
<td>Fed-batch</td>
<td>(Strotbek et al. 2013)</td>
</tr>
<tr>
<td>Small RNA sequencing</td>
<td>CHO-K1</td>
<td>Adherent</td>
<td>Secreted alkaline phosphatase, tissue (SEAP), plasminogen activator (IPA)</td>
<td>Media supplemented with 10% FBS</td>
<td>(Hammond et al. 2012)</td>
</tr>
<tr>
<td>miRNA microarray</td>
<td>CHO-DUX-B11</td>
<td>Suspension</td>
<td>3D6 single-chain Fv-Fc fusion antibody, Human serum albumin (HSA)</td>
<td>Steady state</td>
<td>(Maccani et al. 2014)</td>
</tr>
<tr>
<td>miRNA microarray</td>
<td>CHO-K1, CHO-S</td>
<td>Static</td>
<td>Human IgG1</td>
<td>nil</td>
<td>(Harreither et al. 2015)</td>
</tr>
<tr>
<td>qPCR + Microarray + Quantitative LC-MS/MS</td>
<td>CHO</td>
<td>Suspension</td>
<td>mAb</td>
<td>Batch</td>
<td>(Clarke et al. 2012)</td>
</tr>
<tr>
<td>mRNA &amp; miRNA microarray</td>
<td>CHO-K1</td>
<td>Suspension</td>
<td>IgG1</td>
<td>Batch</td>
<td>(Bort et al. 2012)</td>
</tr>
<tr>
<td>LC-MS/MS</td>
<td>CHO-K1</td>
<td>Suspension</td>
<td>mAb</td>
<td>Fed-batch</td>
<td>(Meleady et al. 2011)</td>
</tr>
<tr>
<td>Small RNA deep sequencing</td>
<td>CHO-K1 (serum-dependent), CHO-k1 (serum –free), CHO-DUXB11 (serum-free)</td>
<td>Static</td>
<td>DUXB11-Epo-Fc, CHO-K1(serum free) -mAb</td>
<td>nil</td>
<td>(Hackl et al. 2011)</td>
</tr>
<tr>
<td>miRNA microarray</td>
<td>CHO-K1, CHO-DUXB11, and CHO-S (among</td>
<td>Static, suspension</td>
<td>Epo-Fc</td>
<td>nil</td>
<td>(Klanert et al. 2016)</td>
</tr>
</tbody>
</table>
1.2.4 MicroRNA-engineering of CHO cells

Mammalian cells are the preferred choice of cell factories for the manufacturing of biologics because of their ability to make proteins with similar posttranslational modifications (PTM) to that of humans unlike bacteria and yeast cells. However, mammalian cell factories or expression systems still grapple with the hurdle of an optimal protein yield. In addition, due to the growing demands of biologics and incremental development of new technologies scientist testing boundaries searching for more ways to optimize mammalian cells most especially CHO cells. Subsequently, different approaches like cellular engineering, process engineering (Alves et al. 2015), media optimization and genetic engineering have applied to CHO cells to optimize productivity. Common trend in genetic approach is the overexpression of a gene of interest which invariably adds more translational burden to the cell machinery. There are growing interest among the scientific community of using miRNA to engineer CHO cells since this approach do not add extra translational burden and the actions of one miRNA could have effect on more than one pathways. Engineering CHO cells with miRNAs have been applied in to improve different cellular processes such as apoptosis (Druz et al. 2013), protein production (Loh, Yang, and Lam 2017; Chen et al. 2005; Fischer, Paul, et al. 2015; Kelly et al. 2015; Meyer et al. 2017; Emmerling et al. 2016; Loh et al. 2014), growth rate (N. Barron et al. 2011)

1.2.4.1 MicroRNA engineering of CHO cells expressing difficult-to-express (DTE) protein

Proteins are said to be difficult-to-express (DTE) if they are insoluble making it difficult for downstream characterization and analyses e.g. Western blotting and purification. Improper
folding (or misfolding) of proteins as well as aggregate formation could be troubling when using producing these proteins especially for commercial purposes. The fusion of human serum albumin (HSA) or Fc to the N-terminus of DTE protein helped in mitigating insolubility (Carter et al. 2010). Co-expression of molecular chaperones (BiP, PDI, CypB), unfolding protein response (UPR) and chemical chaperones (PBA, DMSO, glycerol) have been used to mitigate aggregation problems (Johari et al. 2015). Moreover, in the case of enzymes, most of it could exist in their inactive form. Lastly, expressing a low yield protein is one of the most common problems and scientists are using cellular and process engineering methods to solve this problem. Previous studies have shown that CHO-K1 has a larger ER and higher mitochondrial mass than CHO DUXB11 which might indicate reason why CHO-K1 gives better yield.

Of recent, scientists are using miRNA to engineer CHO cells secreting DTE proteins (e.g. lysosomal proteins, membrane proteins- neurotensin receptor type 1 (NTSR1) and serotonin transporter (SERT) in order to improve relevant bioprocess phenotypes like productivity (Inwood, Betenbaugh, and Shiloach 2018; Pybus et al. 2014; Schoellhorn et al. 2017; Fischer et al. 2017, 2013).

1.3 Difficult-to-express lysosomal enzymes

Lysosomal enzymes are proteins in the lysosome of most eukaryotic cells and are responsible for the breakdown of cell waste products, carbohydrates, lipids, proteins, nucleic acids which are then returned to the cytoplasm. They are produced in the endoplasmic reticulum and modified in the Golgi apparatus. The enzymes are destined for the lysosome since they are tagged with mannose-6-phosphate (m6PO4) which binds with mannose-6-phosphate receptors which directs the enzyme to the lysosome with aid of other mechanisms. Though some of the enzyme may not require m6PO4 receptor for trafficking reason why β-
glucoerebrosidase had its mannose terminated and successfully produced in CHO cells (Desnick and Schuchman 2012). Deficiency in these enzymes can cause the accumulation of their substrate in the lysosome of the cells leading to conditions known as lysosomal storage disorder (LSD).

Since prokaryotic cells do not execute post-translational modification (PTM) of proteins, CHO or other mammalian cells have become the host platform of choice in the manufacturing of lysosomal enzymes. Lysosomal enzymes require PTM for their stability and activity. Production of lysosomal enzyme could be difficult as it has been reported that not all overexpressed recombinant enzyme in a CHO host are successfully secreted into the surrounding culture medium. Consequently, production of a large batch will be required leading to a high cost of production (Ioannou, Bishop, and Desnick 1992; Desnick and Schuchman 2012). Examples of lysosomal enzymes are: glycosidases, protease, sulfatases, acid phosphatases and lipases etc.

**Aim and scope of research**

Though CHO cells have been the commonly used mammalian cell line for the production of recombinant therapeutic protein scientists still believe that the production capacity limit of this workhorse has not been reached. Different cellular, process and genetic engineering means has been used to make this happen. Currently, protein expression scientists are tapping much of the knowledge of miRNA garnered from cancer research and applying that to CHO or other mammalian cell engineering to induce relevant bioprocess phenotypes. One key factor that really promoted this research was the sequencing of the CHO genome, which promoted profiling and cellular engineering of different CHO cell lines. The advantage of using miRNA over gene expression for engineering purpose is that it does not add an extra translational
burden to the machineries of the cell. The scope of this research will rest on the profiling and miRNA-engineering of CHO cell line stably expressing a DTE lysosomal enzyme.

**Significance of research**

The miRNAs and mRNA identified from the next-generation sequencing experiment of CHO cell lines expressing DTE lysosomal enzymes will provide the scientific community potential list of targets for cell engineering for the improvement of cell growth and volumetric productivity. In addition, it provided substantial knowledge of miRNA dynamic changes in CHO cell lines that vary in their growth rate and specific productivity.

Furthermore, identifying sulfatase modifying factor 1 (SUMF1) as a target to miR-23a-3p and miR-377-3p and observing an improvement in sulfatase activity by inhibiting these miRNAs demonstrate the importance of using miRNA to effect desirable bioprocess qualities in CHO cells. This observation could provide a context for research of these miRNAs in clinical research.

**Overview of the study**

This thesis is written in chapters. Chapter 1 has the introduction and literature review. It consist of overview on bioprocessing, microRNA, its origin, mechanisms of action and its application in engineering difficult-to-express proteins. Lastly it contains the aim, scope and significance of the research.

Chapter 2 highlights the hypothesis and objective of the research. Chapter 3 contains full description of the materials and methods. Chapter 4 and Chapter 5 contain miRNA/mRNA profiling research and miRNA engineering experiment respectively.
CHAPTER 2

2.0 RESEARCH OBJECTIVES

Hypothesis

MiRNA has been proven through research to have the ability to regulate major cellular processes in the cells that are relevant to bioprocessing phenotypes of CHO cells as production host. Developing a good understanding of miRNA expression dynamics will help in identifying selected miRNAs as engineering targets whose impact on CHO cell could potentially improve relevant bioprocess phenotype(s).

Research Objectives

1. Identify potential miRNAs as targets for downstream CHO cell engineering by transcriptomic profiling of in house CHO cell stably secreting a difficult-to-express (DTE) enzyme

2. Identify potential miRNAs that targets sulfatase modifying factor 1 (SUMF1) and evaluate their effect on sulfatase specific activity
CHAPTER 3

3.0 MATERIALS AND METHODS

3.1 CELL CULTURE TECHNIQUES AND TRANSFECTION

3.1.1 Biosafety cabinet

Cell culture work were performed in a Logic+ laminar flow biosafety level II cabinet (Labconco, G3371028) to prevent contamination from airborne particulates. Before any cell culture work, the work area inside the cabinet was wiped down with 70% isopropyl alcohol (IPA). All media bottles and cell culture flasks were all wiped down with IPA to prevent transferring contaminants into the cabinet. At the end of all cell culture work, the cabinet was wiped down again with IPA and sash closed.

3.1.2 Cell Thawing

The frozen cells were retrieved from liquid nitrogen and thawed using a liquid water bath (Lab Armor) with gentle swirling or the use of ThawStar thawing system (Biocision). The thawed cells were pipetted and transferred to a T-75 flasks containing 10 mL of media pre-warmed at 37°C. Flask was transferred to a static incubator kept at a temperature of 37°C, 5% CO₂ and 8% humidity for 24 hours. The cells were pelleted by centrifugation at 250 x g for 5 minutes and re-suspended in a fresh media for subsequent culturing in order to remove traces of DMSO contained in cell freezing media.

3.1.3 Cell Freezing

Cells for cryopreservation were counted and harvested when the cell concentration was at about 4.0e6 cells/mL of the exponential cell growth phase. The cells were pelleted by centrifugation at 250 x g for 5 minutes and supernatant discarded. 1 mL of cryopreservation media was added to re-suspend every 10e6 cells pelleted. 1 mL of the mixture containing
10e6 cells was transferred to a cryovial and stored at -80\(^\circ\) C overnight. It was transferred the following day to a liquid nitrogen kept at -180\(^\circ\) C for long term storage.

3.2 Passaging of cell lines

3.2.1 Adherent cells

Cell splitting were done at a ratio of 1:2 when the cells have reached confluency of 80-90%. Before splitting, spent media was removed with a glass pipette or aspirated before rinsing with sterile 1X DPBS (Corning, 21-031CV) to remove any residual phenol contained in the media that will dampen the effect of TrypLE Express (ThermoFisher Scientific, 12605-028), for cell dissociation. After the addition of 2-5 mL TryPLE Express, the flask or plate was transferred to a humidified HERAcell Vios 160i (Thermo Scientific) static incubator kept at 5\% CO\(_2\) and 37\(^\circ\) C for 2-4 minutes. The flask was observed under the Evos XL core microscope (Life Technologies) to make sure all cells have dissociated and assumed a circular morphology. About 3 - 20 mL serum-containing media were added into flask to keep the cells in suspension and pipetted out into a 15 or 50 ml tube depending on the size of the flask. These tubes were then centrifuged at 250 g for 5 minutes. After centrifugation, the supernatant was discarded and the cell pellet re-suspended in a fresh serum-containing media. 20 mL fresh media was used for cells split equally into two T-75 flask (Corning) and 30 ml media for cells split equally into two T-175 flask (Corning).

3.2.2 Suspension cells

Cells to be passaged were first counted on the ViCell™ XR (Beckman Coulter) to obtain viable cell density and viability values. Cells were passaged every 3-4 days at a seeding density of 0.5 x 10\(^6\) cells/mL. The volume of cell culture required to seed cells were aliquoted into a new flask and fresh pre-warmed (37\(^\circ\) C) media was added at the required volume. The flask was
transferred to a stackable incubation shaker (Multitron Pro, Infors HT) maintained at 37° C in atmospheric CO₂ of 5% and 80% humidity for 3-4 days with 125 rpm rotation before passaging.

3.2.3 Cell counting

3.2.3.1 Vicell™ XR cell viability analyzer

500 uL of cell culture from cells grown in flask was aliquoted into a sampling cup and inserted into the ViCell™ XR (Beckman Coulter) sample cup holder for determination of the concentration of cells, viability and cell size. The ViCell uses the Trypan Blue exclusion principle by counting cells stained with dye as dead cells.

3.2.3.2 Nucleocounter NC-200

Cell culture sample from cells grown in 6-well plates was loaded into vial cassette (Chemometec, 941-0011) by dipping the sterile cassette tip into the well containing suspended cells and pressing the piston to aliquot sample (~1.4ul used) into the cassette. The undulating shape of the mixing channel in the cassette allows the sample to mix very well with acridine orange and DAPI dyes contained in this channel. The acridine dye stains all cells, while DAPI stains all non-viable cells. After aliquoting the cell sample, the cassette was inserted into the NC-200 cassette holder and sample analyzed by pressing the “run” button.

3.3 Transfection

3.3.1 Transfection of siRNA/microRNAs with lipid-base reagent method

Transient transfection of suspension cells were carried out by seeding and transfecting the cells on the same day (reverse transfection). For adherent cells, the cells were seeded 24 hours before transfection (forward transfection). Following manufacturer’s protocol,
transfection of suspension cells (e.g. CHO or HEK293 cells), were carried out on the day of transfection by seeding 2.5 mL of cell culture at a concentration of 0.3 - 0.8e6 cells/mL (depending on what cell line was used) in each required well of a 6-well plate. Transfections were carried out in duplicates. All components required for transfection were used in 5-10% excess to compensate for loses due to pipetting. A master mix of microRNA mimic/inhibitor or siRNA was created by diluting 7.5 uL of miRNA mimic/inhibitor or siRNA (20μM) with 485 uL of Opti-MEM I reduced serum media (Gibco, 31985070) in a 5 mL DNA Lobind tube (Eppendorf, 0030108310). The mixture was mixed gently. Lipofectamine™ RNAiMax (Thermo Scientific, 13778150) reagent was mixed gently before use. 7.5 uL of Lipofectamine RNAiMAX were added to the tube containing mixture of miRNA mimic/inhibitor or siRNA and Opti-MEM I reduced serum media for complex formation. The mixture was allowed to incubate for 10-20 minutes at room temperature. After incubation, 500 uL from the mixture was added to each well containing 2.5 mL of 0.3-0.6e6 cells/mL of cells giving a final volume of 3 mL and final concentration of 50 nM. Each 6-well plate was mixed gently by rocking back and forth. Plates were kept in a static incubator for 96 hours.

3.3.2 Transfection of plasmid DNA (pDNA) using electroporation method

Transient transfection was carried out by using Maxcyte STX scalable transfection system (MaxCyte) for adherent or suspension cells. Adherent cells were trypsinized with TrypLE Express (Thermo Scientific, 12605-028) and neutralized with cell media and mixture transferred to a 15 mL tube. However, for suspension cells, the cells were transferred directly to a 15 mL tube. With a target cell culture volume of 30 mL, 20e6 cells in suspension were spun at 250 x g for 5 minutes and supernatant discarded. Cell pellet was rinsed in 1.5 mL of electroporation buffer (Maxcyte, B201-100) and pelleted again at 250 x g for 5 minutes. While the cells were being spun, 20 μg of pDNA was aliquoted into a 1.5 mL microcentrifuge tube.
and the OC 400 processing assemblies (PA) were labeled based on experimental design. The cell pellet were dissolved in the right volume of electroporation buffer (400 uL less volume of plasmid used) and mixed with the plasmid by gentle up and down pipetting. The mixture were transferred to a PA in a final volume of 400 uL avoiding the generation of air bubbles and made sure the liquid was in full contact with both sides of the PA. The PA were inserted into the Maxcyte STX system and the cell transfection protocol was selected depending on what cell line was used. After electroporation, the PA was removed from the STX system and the electroporated cells were transferred to an empty 30 mL Ehrlenmeyer flask (Corning). 5ml of required cell culture media was added to the 400ul electroporated cells contained in the 30ml flask and incubated in a static incubator at 37°C for 20 - 30 minutes. After incubation, 24.6 uL of additional media was added to the flask, which brought the cell density to about 0.3 – 0.6e6 cells/mL. Cells were cultured either in a 6-well plate or cell culture flask.

3.3.3 Cell Imaging/Transfection efficiency measurement

Green fluorescent protein (GFP) plasmid and small-interfering RNA (siRNA) Block-iT Alexa Fluor 555 red fluorescent control were used as indicators to assess transfection efficiency in a lipid-mediated or electroporation transfection method. Media from the well of a 6-well plate containing cells transfected with GFP or Alexa Fluor were aspirated and about 1 mL of 1X DPBS was added in order to eliminate any color background generated by the media when viewed under the EVOS FL Auto fluorescence microscope (Life Technologies). The cells transfected with GFP or Red Alexa Fluor were viewed under the fluorescent channel for GFP or Texas Red (RFP) respectively. The images obtained were viewed under 10X, 20X or 40X magnification.
3.4 MOLECULAR BIOLOGY TECHNIQUES

3.4.1 Molecular biology water

Water for preparing solutions used for assays were made using ultrapure water. The ultrapure was generated from Milli-Q-Advantage A10 water purification system (Millipore). Hypure molecular biology grade water (GE, SH30538.01) was used for reconstituting or rehydrating antibodies and oligonucleotides.

3.4.2 Sterilization

Glassware used for solution preparation were washed and cleaned by laboratory operation unit at BioMarin. Other materials for cell culture and molecular biology were washed with Tergazyme (Alconox), a detergent containing protease enzyme, and rinsed with tap water. Microcentrifuge tubes placed in a cylindrical plastic container with lid were sterilized by autoclaving at 121⁰ C for 45 minutes under a pressure of 1 bar by lab operation personnel.

3.4.3 Reconstitution of lyophilized miRNA

Lyophilized MicroRNAs were reconstituted by the addition of nuclease-free molecular water (GE healthcare LifeSciences, SH30538.01) to a final stock concentration of 20 or 100 μM. After the addition of nuclease-free water, mixing was done by up and down pipetting. The mixture was allowed to incubate for 10 minutes at room temperature before being transferred to -20⁰ C for storage.

3.4.4 Rehydration of antibodies

Primary and secondary antibodies for Western Blotting were rehydrated by the addition of certain amount 1X DPBS to a concentration recommended by the supplier of the antibodies.
After addition of 1X DPBS, mixing was done by up and down pipetting followed by centrifugation at 250 x g for 30 seconds.

3.4.5 Extraction of total RNA (including small RNA - microRNA)

Total RNA, including small RNA, was extracted using Quick-RNA Microprep (Zymo Research, R1050). Samples of suspension cells were pelleted at 250 x g for 5 minutes in a minicentrifuge (Eppendorf) and supernatant discarded. For adherent cells, the cells were dissociated with TrypLE Express (Thermo, 12605028) and pelleted as described above. For both adherent and suspension cells, the cell pellets were resuspended in 200 – 400 uL of RNA lysis buffer and vortexed briefly for homogenization. Particulates were removed by centrifuging the lysates at 14000 x g for 1 minute. To purify the RNA, the supernatant was transferred to an RNase-free tube and 100% ethanol added to the lysis buffer containing the sample at a ratio of 1:1. Reaction mixture was well mixed by up and down pipetting. The mixture was transferred to a Zymo-Spin IC column enclosed in a 2 mL collection tube and centrifuged at 14000 x g for 30 seconds. The flow-through was discarded. For trace DNA removal, the column was washed with 400 uL of RNA wash buffer and centrifuged at 14000 x g for 30 seconds. The flow-through was discarded. DNase I reaction mixture was prepared using DNase I and DNA digestion buffer at a ratio of 1:7 respectively. 40 uL of DNase I reaction mixture was added into the spin column directly, incubated at room temperature for 15 minutes and then centrifuged at 14000 x g for 30 seconds. 400 uL of RNA prep buffer was added to the column, centrifuged at 14000 x g for 30 seconds and flow-through discarded. The column was washed with 700 uL RNA wash buffer, centrifuged at 14000 x g for 30 seconds and flow through discarded. 400 uL of wash buffer was added to the column and spun at 14000 x g for 2 minutes to completely remove residual wash buffer in the column. The column was transferred to a new 1.5 mL tube (Eppendorf) and 40 uL of DNase/RNase-free
water was added into the column and centrifuged at 14000 x g for 30 seconds for RNA elution. The eluted RNA was stored at 80° C.

3.4.6 Determination of nucleic acid concentration

Quantification of double stranded DNA (dsDNA) or single stranded RNA (ssRNA) concentration was carried out using DS-11 spectrophotometer (Denovix) by following manufacturer’s protocol.

3.4.6.1 Plasmid DNA

The dsDNA icon was first tapped. This led to the opening of the dsDNA measurement window. The top and bottom sample surfaces of the DS-11 spectrophotometer (Denovix) was wiped down with a dry lab wipe and 2 μL of nuclease-free water (Qiagen) was pipetted on the sample surface as blank measurement. The top arm was lowered and the “blank” button tapped. After blank measurement, the sample surface was wiped down with fresh dry lab wipe. 2 μL of test sample (dsDNA) solution was pipetted onto sample surface, sample name entered, arm lowered and the “measure” button tapped to obtain absorbance value. The measurement generates DNA concentration in ng/μL including the 260/280 and 260/230 absorbance values. The result was exported through the instrument email messaging system.

3.4.6.2 Total RNA

The same protocol was followed as illustrated above for dsDNA except that RNA icon was tapped for measurement of RNA absorbance values.

3.4.7 Determination of RNA quantity and quality using Bioanalyzer

Total RNA concentration and quality were determined on the Agilent 2100 Bioanalyzer instrument using the RNA 6000 Nano Kit (Agilent, 5067-1511) according to the instruction of
the manufacturer. During sample preparation, total RNA isolated was diluted to 100 ng/µL because the quantitative range for analysis on the instrument falls between 25 – 500 ng/µL. Before each measurement on the Agilent 2100 Bioanalyzer instrument, RNaseZAP (Ambion, 9780) and RNase-free water was used for routine decontamination and cleaning of the electrode respectively. In preparing the gel, 550 µL of RNA gel matrix was pipetted into a spin filter and centrifuged at 1700 x g for 10 minutes at room temperature, aliquoting 65 µL into a microcentrifuge tube. The gel-dye matrix was prepared by equilibrating the RNA dye concentrate to room temperature for 30 minutes. The dye was vortexted for 10 seconds, spun down and 1 µL of the dye was added to the tube containing 65 µL aliquot of filtered gel. The mixture was vortexted and spun in a minicentrifuge at 13000 x g for 10 minutes at room temperature. After placing the RNA chip on the chip priming station (CPS), 9 µL of the gel-dye mix was pipetted into a well marked. Ensuring that the plunger was positioned at 1 mL level, the CPS was closed, plunger pressed down until it was held by the clip and waited for 30 seconds before releasing the clip. After 5 seconds, the plunger was pulled to the 1 mL position. The CPS was opened and 9 µL of the gel-dye mix was pipetted into wells marked. The unused gel-dye mix was discarded. 5 µL of RNA marker was pipetted into all 12 sample wells and in the well marked with a ladder sign. In addition, 1 µL of each RNA sample was pipetted into the sample wells depending on the number of RNA sample being analyzed. 1 µL of the RNA marker was pipetted into each unused sample well. The chip was placed unto an IKA vortex mixer horizontally and vortexted at 2400 rpm for 1 minute. Finally, within 5 minutes the chip was ran on the Agilent 2100 Bioanalyzer instrument to obtain the RNA concentration and RNA integrity number (RIN). The RIN gives an estimate of the quality of the RNA and it ranges from 1 – 10, with 10 being the best quality.
3.4.8 Reverse transcription (cDNA generation)

Reverse transcription experiment was carried out using miScript II RT Kit (Qiagen, 218160) for the generation of complementary DNA (cDNA) following instructions from the manufacturer. Total RNA, containing miRNA, was used as a template. RNA sample stored in -80\(^\circ\) C was thawed in ice, while 10X miScript Nucleic Mix and 5X miScript HiSpec Buffer were thawed at room temperature and later placed on ice. In order to synthesize first-strand cDNA, reverse-transcription master mix was prepared in a microcentrifuge tube for each RNA sample by adding 4 µL HiSpec buffer, 2 µL Nucleic mix, 2 µL miScript Reverse Transcriptase Mix, RNase-free water (amount depends on RNA template amount used) and finally RNA sample (amount used depends on RNA sample concentration used which falls at the upper limit of 10 ng – 2 µg) for total mixture of 20 µL. The master mix was mixed gently, centrifuged and placed on ice. The tube containing master mix was incubated in a thermal cycler (Life Technologies) at 37\(^\circ\) C for 60 minutes, then at 95\(^\circ\) C for 5 minutes in order to inactivate miScript Reverse Transcriptase Mix. Tube was later placed on ice. The 20 µL reaction mixture was diluted by adding 200 µL of RNase-free water, mixed gently and centrifuged. This ensured that cDNA generated fell between 50 pg – 3 ng per PCR. The diluted cDNA was either used immediately for real-time PCR or the undiluted cDNA was stored at -20\(^\circ\) C for future real-time PCR experiment.

3.4.9 Real Time-PCR (qPCR)

The real-time PCR (qPCR) was carried out using a miScript SYBR Green PCR Kit (Qiagen, 218073) with miScript Primer Assay according to the protocol of the manufacturer for the detection of mature miRNA. All the content of each reagent: 2X QuantiTect SYBR Green PCR Master Mix, 10X miScript Universal Reverse Primer, and diluted cDNA was thawed, mixed and centrifuged before use. A reaction master mix was prepared according to Table 1.
Table 3: Reaction master mix for the detection of matured miRNA in a 384-well plate

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume/reaction (384-well)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2x QuantiTect SYBR Green PCR Master Mix</td>
<td>5 µL</td>
</tr>
<tr>
<td>10x miScript Universal Reverse Primer</td>
<td>1 µL</td>
</tr>
<tr>
<td>10x miScript Primer Assay (U6 snoRNA, GAPDH, miRNA of interest)</td>
<td>1 µL</td>
</tr>
<tr>
<td>RNase-free water</td>
<td>2 µL</td>
</tr>
<tr>
<td>cDNA Template (not added in master mix)</td>
<td>1 µL</td>
</tr>
<tr>
<td><strong>Total volume</strong></td>
<td><strong>10 µL</strong></td>
</tr>
</tbody>
</table>

The GAPDH and U6 SnoRNA forward primers were used for normalization of gene expression and mature miRNA expression levels respectively. The reaction master mix was prepared for three technical replicates with extra 10% reaction component used to compensate for loses due to pipetting. 1 µL of the cDNA template was dispensed into individual wells of the PCR plate for each replicate. The reaction master mix was mixed thoroughly and 9 µL dispensed into wells containing cDNA. The plate was tightly sealed with a film and centrifuged at 1000 x g for 1 minute at room temperature. The real-time LightCycler 480 (Roche Diagnostics) was programed according to the shown below. In addition, melting curve analysis was done to verify specificity (primer-dimer).

Table 4: Roche LightCycler 480 cycling conditions

<table>
<thead>
<tr>
<th>Step</th>
<th>Time</th>
<th>Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Activation step</strong></td>
<td>15 min</td>
<td>95⁰ C</td>
</tr>
</tbody>
</table>
The plate was placed in the real-time cycler and program started. The \( C(T) \) method (Schmittgen and Livak 2008) was used to calculate relative expression differences.

### 3.4.10 Molecular cloning

#### 3.4.10.1 Agarose gel preparation

In order to make 500 mL of agarose gel, 5g of Seakem LE Agarose (Lonza, 50004) was measured and poured into an amber bottle followed by the addition of 300 mL of TAE 1X buffer. The amber bottle with mixture was heated up in a microwave for at least 2 minutes to dissolve the agarose powder. 50 \( \mu \)L of 1X of Biotium gel green nucleic acid stain was added into bottle and mixed by gentle swirling. The bottle containing freshly prepared agarose gel was kept at 65\(^\circ\)C in a bead bath (Lab Amor).

#### 3.4.10.2 Gel extraction

This process was carried out using the QIAquick Gel Extraction Kit (Qiagen, 28704) by following the procedures stated by the manufacturer. DNA was purified from small piece of gel that contained our desired DNA. Gel slice containing desired DNA was cut from agarose gel with a scalpel and transferred into a 1.5 mL microcentrifuge tube. The gel slice was weighed to accommodate the maximum amount (400 mg) the spin column could hold. Buffer QG was added to the tube with gel at a volume of 3:1 (100 mg gel \( \sim \) 100 \( \mu \)L). Gel was
incubated at 60°C and vortexed every 2 minutes until gel slice dissolved completely. The gel sample was transferred to a MinElute spin column placed inside a 2 mL collection tube and centrifuged at 17900 x g for 1 minute at room temperature. The flow-through was discarded and process repeated if sample was greater than 800 μL. 500 μL of Buffer QG was added to the column and centrifuged at 17900 x g for 1 minute at room temperature. The flow-through was discarded and column placed back into the collection tube. Similar process as above was repeated using 750 μL of Buffer PE; however, the column was allowed to incubate for 3 minute after adding Buffer PE if DNA was intended for sequencing or blunt-ended ligation downstream. After discarding the flow-through to completely remove residual ethanol, the column contained in 2 mL collection tube was centrifuged at 17900 x g for 1 minute at room temperature. Column was placed in a new 1.5 mL microcentrifuge tube and 10 μL Buffer EB added to the center of the column membrane. To elute the DNA, the column was allowed to stand for 1 minute and centrifuged at 17900 x g for 1 minute at room temperature.

### 3.4.10.3 PCR purification

For the purification of PCR products, QIAquick PCR Purification Kit (Qiagen, 28104) was used according to specifications by the manufacturer. 5 volumes of Buffer PB was added to 1 volume of the PCR reaction and mixed gently. The sample was transferred to a QIAquick column placed in a 2 mL collection tube and centrifuged at 17900 x g for 60 seconds at room temperature. After discarding the flow-through, 750 μL of wash Buffer PE was added and centrifuged at 17900 x g for 60 seconds at room temperature. In order to remove residual wash buffer, the column in 2 mL collection tube was centrifuged again for 1 minute at same conditions. After placing the column in a new 1.5 mL microcentrifuge tube, 30 μL of Buffer EB was added to the center of column, allowed to stand for 1 minute and centrifuged at 17900 x g for 1 minute at room temperature in order to elute DNA.


3.4.10.4 Plasmid DNA purification (maxiprep)

Automated maxiprep purification of plasmid DNA was carried out using the BenchPro 2100 Maxicard Plasmid Purification Instrument (Life Technologies) according to the purification protocols of the manufacturer. The instrument and air compressor was turned on and made sure that the air compressor wasn't letting out air. Setup button was tapped to set up the instrument for maxiprep purification. The reagent tray foil was pierced with a piercing tool and an empty 2 mL collection tube with cap removed was placed into the elution tube slot. After opening the drawer of the instrument, the waste tray was removed and reagent tray placed in the waste tray slot ensuring the tray aligned with the slot. A cell liner was placed in the waste tray reservoir slot. The drawer was opened and waste tray containing reagent tray was placed into the waste tray slot. About 125 mL of *E. coli* culture was poured into the cell liner contained in the reservoir and covered with a lid. The drawer was closed gently which was followed with the insertion of the Maxiprep Card into a slot in the BenchPro 2100 instrument. The start button was tapped, following by the click of the run button in order to run the protocol. At the completion of the run, the card was removed from the slot and discarded. The drawer unit was opened to remove the elution tube containing purified plasmid DNA. The tube was capped and stored at -20°C. The reagent tray and cell liner were discarded appropriately and the waste tray decontaminated with bleach and washed with detergent for future use.

3.4.10.5 Transformation

2 µL ligation reaction was added into vial containing One Shot TOP10 chemically competent *E. coli* and gently mixed. This was incubated on ice for 20 minutes. The cells in the vial was heat-shocked at 42°C for 30 seconds and transferred immediately to ice. 250 µL of super optimal broth with catabolite repression (S.O.C) medium kept at room temperature was added into each vial and incubated at 37°C for 1 hour with shaking. Depending on the size of the
plasmid, 20 – 100 µL of bacterial culture was dropped using a pipette close to the edge of the LB agar plate containing 50 µg/mL spectinomycin. An inoculation loop was used to streak the dropped bacteria culture on the LB agar plate in a clockwise manner. The plate was incubated overnight at 37°C.

3.4.10.6 Purification of plasmid DNA (miniprep)

QuickLyse Miniprep Kit (Qiagen, 27505) was used by following the protocols of the manufacturer. It was ensured that the OD$_{600}$ of E. coli cells from 2 mL culture fell between 1.8 – 2.2 using a nano-spectrophotometer (Denovix). If it is greater than 2.2, the culture was diluted to 2.0 with deionized DNA-free water. 1.5 mL of the bacteria culture was pipetted into 2 mL QuickLyse Lysis Tube and centrifuged at 18000 x g for 1 minute at room temperature. The supernatant was decanted by inverting the tube on a paper towel. 400 µL of ice-cold Complete Lysis Solution (placed on ice) was added to bacteria pellet and vortexed at the highest setting for 30 seconds. The vortexing was continued until cell clumps were completely resuspended. The lysate was incubated at room temperature for 3 minutes and pipetted into a QuickLyse spin column. The spin column was centrifuged at 17000 x g for 60 seconds. The flow-through was decanted and spin column placed back into tube. The spin column was washed by adding 400 µL of Buffer QLW and centrifuged at 17000 x g for 60 seconds. The flow-through was discarded. The spin column was placed in a new waste tube and centrifuged at 17000 x g for 1 minute to dry the spin column. After transferring the spin column into a new 1.5 mL microcentrifuge tube, 50 µL Buffer QLE was pipetted into the center of the spin column to elute DNA by centrifuging at 17000 x g for 60 seconds.

2.4.10.7 Restriction endonuclease digest

Restriction enzyme digest carried out was adapted from protocol recommended by New England BioLabs (NEB). The purpose of the experiment was to confirm the quality (clean
bands and right sizes) of the plasmid purified using the miniprep (Qiagen). Total reaction volume was 50 µL for digestion of 1 µg of substrate. In order to identify the right enzyme for digestion, the plasmid vector map was viewed on Snapgene software. After identifying at least two enzymes, 10U of each enzyme was used for digestion without exceeding 10% of total reaction volume (50 µL). Typically, after determining the right volume to be used for each reagent, HyPure deionized water (GE Healthcare) was first added into a PCR tube to make up the remaining difference of the 50 µL of total reaction volume. Next, 5 µL of Cutsmart buffer (NEB) was added into the PCR tube. The type of buffer used depended on the enzyme used. The enzymes are better suited to different buffers. This was followed by the addition of 1 µg of the plasmid DNA and the right volume of enzymes were added lastly. The PCR tube containing reaction volume was incubated in Veriti 96-well thermal cycler (Applied Biosystems) at 37° C for 1 hour. While the reaction was incubating, 1% agarose gel was poured into an electrophoresis chamber which was allowed to solidify for about 20 - 30 minutes. 1X TAE buffer was poured into the electrophoresis chamber containing agarose gel to a point a little above the gel level. After the samples have finished incubating, 10 µL of 6X gel loading dye (New England BioLabs, B7024S) was added to all the samples. 15 µL of each sample were loaded into each lane and 10 µL DNA ladder (ThermoFisher, SM0312) in one lane. The gel was ran for about 45 minutes – 1 hour at 100V. At the end, the gel was imaged (Azure Biosystems, c200) and compared to the predicted image as depicted by the Snapgene software gel electrophoresis predictive tool when simulated with the restriction enzymes used.

3.4.10.8 BigDye® Polymerase Chain Reaction (PCR) cycle sequencing

Purified DNA plasmid was amplified using BigDye® cycle sequencing (Applied Biosystems) protocol as recommended by the manufacturer. The table below shows the amount of reagent added for each reaction in a microcentrifuge tube.
Table 5: BigDye reaction mix preparation

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. BigDye ready reaction mix</td>
<td>8 µL</td>
</tr>
<tr>
<td>2. Plasmid DNA template (500 – 1000 ng)</td>
<td>2 µL</td>
</tr>
<tr>
<td>3. Primer @ 3.2 pmol/ µL (Forward &amp; Reverse)</td>
<td>1 µL</td>
</tr>
<tr>
<td>4. Deionized water</td>
<td>9 µL</td>
</tr>
<tr>
<td><strong>Total volume</strong></td>
<td><strong>20 µL</strong></td>
</tr>
</tbody>
</table>

Each reaction was carried out separately for the forward and reverse primer of the DNA template. After adding all reagents in the order listed in Table 3, each tube was mixed well and spun briefly. The tubes were placed in Veriti thermal cycler (Applied Biosystems), volume of reaction set to 20 µL and PCR cycle parameters entered as seen in Table 4 below and PCR reaction started.

Table 6: BigDye PCR reaction parameters

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>Duration</th>
<th>Cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td>96</td>
<td>1 min</td>
<td>1</td>
</tr>
<tr>
<td>96</td>
<td>10 sec</td>
<td></td>
</tr>
<tr>
<td>50 (depending on template primer)</td>
<td>5 sec</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>4 min</td>
<td>25 cycles</td>
</tr>
<tr>
<td>4</td>
<td>∞</td>
<td>hold</td>
</tr>
</tbody>
</table>

3.4.10.9 DNA sequencing clean-up

This process was carried out using ZR DNA sequencing clean-up kit (Zymo Research, D4050) according to the protocol of the manufacturer. 240 µL of sequencing binding buffer was added
to the 20 µL PCR sequencing reaction sample and mixed. The mixture was transferred into Zymo-Spin IB Column in a collection tube. The spin column with collection tube was centrifuged at 10000 rpm for 1 minute. 300 µL of wash buffer was added to the column and centrifuged at 10000 rpm for 1 minute. Spin column was placed into a new 1.5 mL microcentrifuge tube. 20 µL of Hi-Di formamide solution was added into the column and centrifuged at 10000 rpm for 1 minute. The eluted PCR product was transferred to a clean Axygen 96-well sequencing plate for subsequent use with the sequencing instrument.

3.5 PROTEIN ANALYSIS

3.5.1 Bicinchoninic acid (BCA) Protein Assay (microplate procedure)

A stock of the diluted standards (Bovine serum albumin-BSA) was prepared in clean Eppendorf tubes with concentration ranging from 2000 µg/mL to 25 µg/mL and using the diluent (distilled water) as blank. The cell lysate to be tested was diluted to bring total protein concentration to between 2000 µg/mL and 25 µg/mL based on historical or previous experience with cell type and cell culture condition. 10 µL of each standard and sample in duplicates were added to the corner of the well in a clear 96-well plate (Corning, 3641). Finally, 200 µL of the working reagent (WR) was added into all wells and mixed well by up and down pipetting. The microplate was transferred to a Jitterbug (Boekel Scientific, 130000-115V), a heated microplate shaker, mixed for 10 minutes and incubated at 37° C for 30 minutes. Thereafter, absorbance reading was measured at 562nm on a SoftMax plate reader.

3.5.2 Western Blot (Immunoblotting)

The harvested cells were pelleted by centrifugation (Eppendorf, centrifuge 5424) at 2500 rcf for 10 minutes and supernatant discarded. MPER Mammalian Protein Extraction Reagent Kit (Thermofisher, 78501) and Halt Protease Inhibitor cocktail (Thermo Scientific, 87785) were
mixed at a ratio of 100:1 respectively. A certain amount of volume (depending on the cell density) of the mixture, between 100 – 400 uL was added to the pelleted cells. This was mixed by pipetting up and down. The mixture was shaken gently for about 10 minutes and centrifuged at 14000 rcf for 15 minutes to remove cell debris. Supernatants were transferred into a new 1.5 mL microcentrifuge tube (Eppendorf) for storage at 80⁰ C or used immediately for sample preparation. The samples were prepared with a mix of sample buffer, reducing agent and distilled water using manufacturer protocol. The samples were then heated to 95⁰ C in a microcentrifuge heating block (VWR) for 10 minutes. Samples were loaded into wells in a 4 - 12% SDS-polyacrylamide gel Nupage premade gels (Novex, Invitrogen by Thermo Scientific) containing 1X MOPS SDS running buffer (Novel by Life Technologies, NP0001). 10 uL of SeeBluePlus2 pre-stained protein ladder (Thermo Scientific, LC5925) was used as a molecular weight standard. Required volume of samples depending on gel size used was added to the remaining wells. Gel was ran in a PowerPac 200/300 electrophoresis power supply (Biorad) at 100V for 5-10 minutes, then increased to 110V after samples were out of the wells and into the stacking gel. Samples were allowed to run until they get to the bottom of the gel. The stacked gel casing was rinsed with deionized water. Proteins were transferred from the gel to a nitrocellulose membrane using the iBlot 2 dry blotting system (Life Technologies) with iBlot 2 transfer stacks (Invitrogen, IB23002). The membranes were allowed to wash in a blocking buffer (Thermo Scientific, 37543) for 1 hour on a shaker at room temperature. The following primary antibodies: human SUMF1 (1:1000 #TA337720 Origene), mouse/human SUMF1 (1:500 #MAB2779 R&D Systems), GAPDH (1:5000 #PA1-988 Thermo Scientific), were diluted in blocking buffer and used to probe the membrane overnight at 4⁰ C on a shaker. Next, the membranes were probed with fluorochrome-conjugated secondary antibodies and detected with Odyssey CLx Imaging System (LICOR Biosciences) after 1 hour of incubation.
CHAPTER 4

mRNA and miRNA profiling of a high and low CHO cell producer clones stably expressing a difficult-to-express lysosomal protein using next generation sequencing

Ifeanyi Michael Amadi¹, ², Kevin Ru¹, Vishal Agrawal¹, Terri Christianson¹, Karen Yu¹, Cameron Bardliving², Parviz Shamlou², Jonathan Lebowitz¹

¹ BioMarin Pharmaceutical Inc. 105 Digital Dr. Novato, California 94949, USA
² Keck Graduate Institute. 535 Watson Dr. Claremont, California 91711, USA

ABSTRACT

Chinese hamster ovary (CHO) cells have become the workhorse for the manufacturing of recombinant therapeutic protein and the use of a fed-batch biphasic (temperature shift at the early part of cell growth) process have tremendously improved the volumetric yield of
most well understood therapeutic proteins. However, difficult-to-express (DTE) recombinant therapeutic protein like complex multi-specific proteins, DTE mAbs, and lysosomal enzymes have seen difficulties in manufacturability of CHO cells and other mammalian cells as production platforms. In addition, genetic heterogeneity and clonal variations in CHO cell lines have been shown to produce cells with variable cellular growth rate, product yield and specific productivity. In this study, we sought to understand the gene and miRNA expression profile dynamics of three in-house CHO cell lines (non-producer, low and high producers) stably secreting a DTE lysosomal protein but differ in their growth rate, specific productivity and titer. MicroRNA are short non-coding RNA with 18-24 nucleotide in length. They have been shown to play major roles in various cellular processes with relevance and impact in bioprocessing phenotypes. Next-generation sequencing (NGS) technology was employed for mRNA and miRNA expression analysis of samples taken on day 3, 5, 7 and 10 from CHO cells grown in a 5L fed-batch biphasic bioreactor. The RNA-Seq and small RNA-Seq data were analyzed by the Bioinformatics group at BioMarin. Correlation and differential analyses from the analyzed data identified 500 mRNAs and 35 miRNAs that were differentially expressed (DE) in the in-house CHO cell lines over cultivation period. Meanwhile, from the 35 DE miRNAs, cgr-miR-31-5p represent the top DE miRNA with an upregulation of 131 fold in the higher producer compared to the lower producer. In addition, correlation results of mRNA and miRNA expression profile were also determined. Taken together, this study has been able to provide miRNA and mRNA expression dynamic changes in CHO clones that differ in growth and productivity profiles, while providing potential targets for CHO cell engineering for desirable bioprocess phenotype.

**Keywords:** Chinese hamster ovary (CHO cells), cell engineering, microRNA (miRNA), lysosomal protein, Next generation sequencing (NGS), difficult-to-express (DTE)
**Abbreviations:** CHO, Chinese hamster ovary; DTE, difficult-to-express; miRNA, microRNA; UTR; qRT-PCR, real-time quantitative reverse-transcription PCR; CHO-lys, CHO cell secreting lysosomal protein
4.0 Introduction

Among mammalian cells used for recombinant therapeutic production, CHO cell has remained the dominant player as most approved biologics were manufactured in CHO cells. Moreover, tremendous efforts by scientists have helped in pushing the productive capacity of CHO cells to a much higher volumetric yield via media optimization (Buchsteiner et al. 2018; Clincke et al. 2011), cellular (Dorai et al. 2010; Druz et al. 2013), process (Gagnon et al. 2011) and genetic engineering (Rajendra, Peery, and Barnard 2016; Jazayeri et al. 2018). However, the latter approach entails the manipulation of gene expression which could add an extra translational burden on the host cell (Migani, Smales, and Bracewell, 2017). Recently, microRNA-engineering of CHO cells have shown proven efficacy at improving cell specific productivity (Fischer, Paul, et al. 2015; Kelly et al. 2015; Meyer et al. 2017; Emmerling et al. 2016; Loh et al. 2014), growth rate (Strotbek et al. 2013), and apoptosis resistance (Druz et al. 2013) in CHO cell lines expressing less complex and well understood recombinant therapeutic proteins.

MicroRNA (miRNA) are small (18-24 nucleotide) non-coding RNA with the ability to regulate ~30% of gene expression in cells post-transcriptionally (Pasquinelli 2012; Bratkovič et al. 2012; Winter et al. 2009; Ha and Kim 2014). The pleiotropic function of most miRNA allows one miRNA to target more than one mRNA transcripts (Fischer, Handrick, et al. 2015). MiRNA transcription begins at the nucleus and is transcribed by RNA polymerase II into a long primary transcript (pri-miRNA) with multiple stem loops. Thereafter, the pri-miRNA is processed in the nucleus by microprocessors (Drosha and DGCR8) into a hairpin structure (~72 nt long) called precursor miRNA (pre-miRNA) and exported via exportin-5 (XPO5) into the cytoplasm in a RanGTP-dependent manner. While in the cytoplasm, Dicer, an endoribonuclease, trims the stem loop of the pre-miRNA forming a miRNA duplex (~22 nt) which consist of 5p or 3p mature strands. Either the 5p or 3p strand is selected (the other is
either degraded, but not in all cases) for attachment with Argonaute protein family to form a complex called miRNA-induced silencing complex (miRISC). The miRNA directs the miRISC to a potential mRNA transcript target for inactivation via translational repression, deadenylation or degradation (O’Brien et al. 2018; Pasquinelli 2012). A couple of studies have used next generation sequencing (NGS) (Hackl et al. 2011; Stiefel et al. 2016; Monger et al. 2015; Hammond et al. 2012), microarray technology (Maccani et al. 2014; Klanert et al. 2016; Bort et al. 2012; Harreither et al. 2015), mass spectrometry (Clarke et al. 2012; Meleady et al. 2011) and microRNA library (Strotbek et al. 2013) to profile the transcriptome (miRNA and/or mRNA) or proteome in CHO cells cultivated either in batch (Clarke et al. 2012; Bort et al. 2012; Harreither et al. 2015), fed-batch (Stiefel et al. 2016; Klanert et al. 2016; Meleady et al. 2011), steady-state mode (Maccani et al. 2014) or in adherent cell culture (Hammond et al. 2012; Hackl et al. 2011). However, only few of these studies have used CHO cell lines secreting a difficult-to-express (DTE) protein. Studies that have evaluated mRNA and/or miRNA expression profile changes in CHO cells secreting DTE protein (e.g. EPO-Fc fusion protein) either used microarray technology for analysis or steady-state cultivation of the CHO cell lines (Maccani et al. 2014; Klanert et al. 2016); microarray technology doesn’t have the high sensitivity of NGS for low (or rare) and high gene abundance (Hurd and Nelson 2009) just as steady-state cultivation may not give a true picture of miRNA expression changes as in a traditional fed-batch process commonly used in the industry. Currently, few works have be done testing the limit of miRNA-engineering of CHO cells secreting DTE proteins such as complex multi-specific proteins, DTE mAbs. However, most DTE proteins like lysosomal enzymes have seen difficulties in manufacturability using CHO cells or other mammalian cells as production platforms (Migani, Smales, and Bracewell 2017; Liu et al. 2017).
In this study, we sought to understand transcriptional (mRNA and miRNA) expression changes in CHO cell lines (CHO-lys-low producer, CHO-lys-high producer, and CHO-non-producer) secreting a DTE lysosomal enzyme and cultivated in a fed-batch biphasic mode within a 5L bioreactor. 500 DE mRNA (FDR<0.05 & FC>2 or <0.5) and 35 DE miRNA (FDR<0.05) were identified. Among the 35 DE miRNAs, cgr-miR-31-5p was found to be 131 fold upregulated in CHO-lys-high producer compared to CHO-lys-low producer. In addition, miRNA associated with growth rate and titer were also determined.

4.1 Materials and Methods

4.1.1 Cell lines and cell culture

CHO suspension cell line (CHO-non-producer) and two other in house producer cell lines derived as clones from the CHO-non-producer stably express a lysosomal protein and were cultivated in CD CHO medium (Thermo Scientific, Waltham, MA, USA), but supplemented with 8 mM GlutaMax (Thermo Scientific, Waltham, MA, USA) for CHO-non-producer only. The producer cell lines differ by their specific productivity (also in titer and growth rate), so are identified as low and higher producer relative to their production capacity. The high producer (CHO-lys-high) have a specific productivity and titer that is 3x and 2x higher respectively compared to the low producer (CHO-lys-low). However, CHO-lys-low shows a growth rate that is 2x that of the CHO-lys-high producer. Inoculation cells were sourced from cells maintained with a working volume of 30-150 mL within a shake flask (Corning, Oneonta, NY, USA) at 37°C, 5% CO2 and 80% humidity in an orbital shaker incubator (Infors HT, Annapolis Junction, Maryland, USA) agitated at 125 rpm. Cells were passaged every 3-4 days at a seeding density of 0.5 x 10^6 cells/mL. Cell concentration and viability were measured using the ViCell™ (Beckman Coulter, Brea, CA, USA) by the trypan blue exclusion method. For functional validation study of target miRNAs, CHO suspension cells
were cultivated in CD CHO medium ((Thermo Scientific, Waltham, MA, USA) at a seeding density of 0.9-1 x 10^6 cells/mL and propagated in a 6-well plate for 72 hours (Life Technologies, Grand Island, New York). Culture was maintained at 37°C, 5% CO2 and 80% humidity in a static incubator (Thermo Scientific, Waltham, MA, USA) and samples taken at 48 h and 72 h for RNA isolation and protein concentration/western blotting analysis respectively.

4.1.2 Bioreactor fed-batch cultivation

CHO-non-producer, CHO-lys-low, and CHO-lys-high producer cell lines were cultivated in 5L Dasgip bioreactors in a biphasic (temperature shift on D5, from 37°C to 30°C) fed-batch mode. The bioreactors were inoculated at a seeding density of ~0.4 x 10^6 cells/mL with a working volume of 3L and cultivated for 14 days. The cells were cultured in a CD CHO medium (Thermo Scientific, Waltham, MA, USA) and maintained at 37°C, pH 6.9 (Figure 3D), dissolved oxygen 30% and 80 rpm agitation. On day 3 and 5, 5% and 10% efficient feed B media was added respectively. In addition, to maintain adequate supply of energy source for cells to growth and synthesize products, glucose levels were maintained at 2 g/L. Moreover, offline monitoring of cell health, substrate, and metabolic concentrations were measured using the ViCell™ (Beckman Coulter, Brea, CA, USA) by the trypan blue exclusion method for cell concentration and viability, while the Bioprofile Flex Analyzer (Nova Biomedical, Waltham, MA, USA) was used for glucose, lactate, glutamine, and glutamate concentration measurements.
4.1.3 Transient transfection of mimics and inhibitors

Transient transfection of miRNA mimics (mmu-miR-31-5p) and inhibitors (anti-mmu-miR-31-5p) (Qiagen, Germantown, MD, USA; Dharmacon, GE Healthcare, Lafayette, CO, USA) were carried out using Lipofectamine™ RNAiMax (Thermo Scientific, Waltham, MA, USA) transfection reagent and Opti-MEM I (Thermo Scientific, Waltham, MA, USA) reduced serum medium for complex formation at a final concentration of 50 nM, unless indicated differently. The scramble universal negative control (Sigma-Aldrich, Louis, MO, USA) was used as a non-targeting control (NT-siRNA).

4.1.4 Real-time Quantitative Reverse-Transcription PCR (qRT-PCR)

Briefly, cDNA was synthesized from 1 μg of total RNA using miScript II kit (Qiagen, Germantown, MD, USA) for quantification of mature miRNA(s) expression and Superscript Vilo kit (Invitrogen, Thermo Scientific, Carlsbad, CA, USA) for gene expression quantification according to the protocol of the manufacturers. For the quantification of matured miRNAs and gene expression, qRT-PCR was performed with 10⁻¹ diluted cDNA on a LightCycler 480 (Roche Molecular Diagnostics, Pleasanton, CA, USA) using Ssofast™ Evagreen® Supermix (BioRad, Hercules, CA, USA) according to the protocols of the manufacturer. GAPDH and U6 snoRNA were used for the normalization of gene expression and mature miRNA expression levels respectively. miScript Universal Reverse Primer (Qiagen, Germantown, MD, USA) was used for mature miRNA analysis. Details of primers used for qRT-PCR (Table 14). For the quantification of relative gene expression, the comparative C (T) method was employed.
4.1.5 **Western blot analysis**

Protein extraction on harvested cells was done using the M-PER mammalian Protein Extraction Reagent Kit (Thermo Scientific, Carlsbad, CA, USA) according to the protocol of the manufacturer. Halt Protease Inhibitor Cocktail (Thermo Scientific, Carlsbad, CA, USA) was mixed with the extraction solution at a ratio of 1:100 respectively. Total protein was determined using the bicinchoninic acid (BCA) protein assay (Thermo Scientific, Carlsbad, CA, USA) according to the protocol of the manufacturer. Cellular proteins were separated in a 4-12% SDS-polyacrylamide Nupage premade gels (Thermo Scientific) using a 1X MOPS SDS running buffer. Proteins were transferred from the gel to a nitrocellulose membrane using the iBlot 2 dry blotting system (Life Technologies). The membranes were allowed to wash in a blocking buffer (Thermo Scientific) for 1 hour on a shaker at room temperature. The primary antibody against the secreted lysosomal protein was in house and sourced in bulk from a biotech company. The other primary antibody used was GAPDH (1:5000, PA1-988, Thermo Scientific). All primary antibodies were diluted in blocking buffer and used to probe membrane overnight at 4°C on a shaker. Thereafter, the membranes were probed with fluorochrome-conjugated secondary antibodies (LICOR Biosciences, Lincoln, Nebraska, USA) and detected with Odyssey CLx Imaging System (LICOR Biosciences, Lincoln, Nebraska, USA) after 1 hour of incubation.

4.1.6 **Enzymatic assay (titer)**

For determination of titer values from sample supernatants, 4-MUGAL was used as substrate and carried out in a 96-well plate format. All standard and in-house control samples were diluted to appropriate concentrations. Substrate mixture were loaded into plate wells, and incubated at 37°C for 40 minutes. Fluorescence was read at 325 nm
(excitation) and 420 nm (emission) on a SpectraMax M3 plate reader (Molecular Devices, San Jose, CA, USA)

4.1.7 Bioinformatics analyses

To determine the differential expression of global gene and miRNA expression profiles in the lower producer (CHO-lys-low), higher producer (CHO-lys-high) and parental cell line (CHO-non-producer) and their correlation with relevant bioprocess phenotypes (e.g. growth, titer) samples were taken on day 3, 5, 7, and 10 and processed using next generation sequencing (NGS) by an external company. The RNA-seq and small RNA-seq data analyses was performed by the Bioinformatics group at BioMarin. For NGS, the parameters used can be found in Table 7. The prospective target genes (integration of predicted and gene target from RNA-seq data) of the 35 DE miRNAs were obtained from the results of the RNA-seq data analyses. Functional annotation of predicted genes was analyzed using Database for Annotation, Visualization and Integrated Discovery (DAVID) online tool (https://david.ncifcrf.gov/).

4.2 Results

4.2.1 Fed-batch cell culture cultivation

Three CHO cell lines were cultivated in a biphasic fed-batch process using 5L Dasgip bioreactors in quadruplicate for each cell line. One of the cell lines is CHO-parental (CHO-non-producer) and the other two are CHO-lys-low and CHO-lys-high producers derived as clones from the parental cell secreting a difficult-to-express (DTE) lysosomal enzyme. CHO-lys-low and CHO-lys-high producer clones were classified as low and high producers respectively based on their relative differences in volumetric and specific productivity (Table 8). Although, CHO-lys-low has a better growth rate than CHO-lys-high (Figure 2A). To
maintain adequate energy source for the growth of the cells, glucose levels was maintained at 2 g/L (Figure 3C) after temperature shift to 30°C on day 5. As the common process in bioprocessing, the bioreactor temperature reduction to 30°C induced a reduction in cell growth from day 6 (Figure 2A-B) but an increase in overall cell specific productivity (qP) and higher viability maintenance (Figure 2C). CHO-lys-high showed a qP that is about three times 3-fold (1.47-2.88 pg/cell/day) higher than that of CHO-lys-low producer (0.54-1.01 pg/cell/day) (Figure 3A). In addition, CHO-lys-high showed a titer that is about 2 fold (~262 mg/mL) higher than that of CHO-lys-low producer expressing a final mean titer of ~154 mg/mL after day 14 harvest (Figure 3B). Based on the productivity differences between the two clones we sought to understand the gene and microRNA profiles changes between these clones, including the parental cell line.

4.2.2 Identification of growth-correlating miRNAs and mRNAs

To link the expression profile of genes and miRNAs to relevant bioprocess phenotype like growth rate and titer, correlation analyses were done by the Bioinfomatics group at BioMarin. The LogFC and FDR<0.05 of the miRNAs and mRNAs were applied for correlation with growth rate. In total, the number of DE genes that were correlated with growth rate in CHO-non-producer, CHO-lys-low producer and CHO-lys-high producer were 6964, 8637, and 8945 respectively. Furthermore, by overlapping all three cell lines for identification of growth-associated genes, 529 positively correlated and 758 negatively correlated genes were uniquely identified in the CHO-lys-high producer cell line. For miRNAs associated with growth rate in the CHO-non-producer, CHO-lys-low producer and CHO-lys-high producer, 62, 81 and 37 miRNAs were identified. By overlapping these three
cell lines, 8 miRNAs (2 positively correlated and 6 negatively correlated) (Table 9) associated with growth were uniquely identified in CHO-lys-high producer.

4.2.3 Identification of titer-correlating miRNAs and mRNAs
Likewise, as described above, miRNAs and mRNAs with association to titer in both CHO-lys-low and CHO-lys-high producers were determined. Overall, 8062 DE genes in CHO-lys-low and 8093 DE genes in CHO-lys-high were determined to be associated with titer. In addition, 1021 positively correlated and 668 negatively correlated genes were uniquely identified in the CHO-lys-high producer cell line. Moreover, 88 and 56 miRNAs were identified to be associated with titer in CHO-lys-low and CHO-lys-high producer cell lines respectively. Overlapping the CHO-lys-low and CHO-lys-high producers, 17 positive correlated miRNAs and 5 negative correlated miRNAs were identified to be uniquely associated with CHO-lys-high producer (Table 10). Validated targets of these 17 positively correlated miRNAs could be seen in Table 15.

4.2.4 Functional confirmation of miRNA expression by qRT-PCR
To confirm NGS results, we conducted qRT-PCR on selected 14 DE (FDR<0.05 & FC>2 or <0.5) miRNAs derived from the 35 DE (FDR<0.05) miRNAs when CHO-lys-low producer was compared to CHO-lys-high producer cell line. The RNA samples tested were taken from day 7 after temperature shift on day 5 for each replicate bioreactors and we confirmed that the expression of the miRNAs selected across the different cell lines do not overlap on this day. Surprisingly, we could only confirm a 15-fold (131 fold From NGS result) increase of cgr-miR-31-5p expression in CHO-lys-high producer compared to CHO-lys-low (Figure 4A-J). Though we observed a 2.43 fold increase in cgr-miR-221-5p expression in CHO-lys-high using qRT-PCR, the result was in contrast to the one fold decrease observed from the NGS
result. In general, we could confirm the upregulation of cgr-miR-31-5p expression in CHO-lys-high producer compared to CHO-lys-low producer.

4.2.5 Functional effect of miR-31-5p on the secreted lysosomal protein of CHO-lys-low and CHO-lys-high producers

Since we observed a 15 fold (qRT-PCR) increase in cgr-miR-31-5p expression in CHO-lys-high producer compared to CHO-lys-low producer, we transiently transfected both cell lines with mmu-miR-31-5p mimics and inhibitors and cultured for 72 h in a static condition. In addition, miR-31-5p overexpression and downregulation were confirmed after 48 h (Figure 5A-B). After 72 h of post-transfection using the mimics and inhibitors, we observed no significant changes in cell viability and growth profile in the transfected cells compared to the non-targeting control (si-NTC) (Figure 6A-B). Surprisingly, overexpression of mmu-miR-31-5p in CHO-lys-high producer significantly increased the titer of the secreted lysosomal protein compared to the scramble (non-targeting control) while downregulation of mmu-miR-31-5p had no significant effect (Figure 6E). In contrast, overexpression of mmu-miR-31-5p decreased the titer of the secreted lysosomal protein of the CHO-lys-low producer (Figure 6D) and an opposite effect was observed when mmu-miR-31-5p expression was downregulated. (Figure 6C).

4.2.6 Identification of miR-31-5p target genes and functional annotation

To identify the target genes of miR-31-5p, we used DAVID in silico tool for gene ontology (GO) functional cluster analysis. This help identify pathways that are associated with a gene or group of genes. This analysis was conducted by using the list of miR-31-5p target genes (integration of predicted (by atleast 3 miRNA predictive tool) and NGS DE gene lists). Moreover, a mouse gene over a mouse genome function in the DAVID tool was used for
the analyses since using a CHO genome background wasn’t comprehensive enough for the number of genes tested. Only the annotated clustered groups (~#12) with adj.P.Val<0.2 (Benjamini or FDR) were selected for downstream analysis (Figure 7A). Genes related to protein binding, endoplasmic reticulum, cytoplasm and membrane were highly expressed. However, we went ahead to streamline the target genes into 9 potential targets by selecting genes with relevant effect on bioprocess phenotypes and those identified in any of the 12 clustered groups by at least twice. These target genes could be tested for possible functional studies to examine the effect of miR-31-5p on the target genes and protein expressions (Table 11 & 12). Furthermore, the expression of the selected genes have been shown in different studies to have an effect on protein production, protein trafficking, and cell proliferation (Table 13).

4.3 Discussion

Different studies have evaluated the impact of miRNA and/or mRNA expression profile dynamics on relevant bioprocess phenotypes of CHO cells (Harreither et al 2015; Monger et al. 2015; Clarke et al. 2012; Bort et al. 2012; Klanert et al. 2016; Stiefel et al. 2016; Maccani et al. 2014). Moreover, technologies like next-generation sequencing (NGS), microarray and mass spectrometry have been used singly or in combination for standalone or integrated analysis of mRNA, miRNA or protein profiles of CHO cells secreting proteins such as mAb, SEAP, Epo-Fc. Moreover, much of these studies have indicated that global miRNA expression profiles is either product and/or clone specific (Stiefel et al. 2016; Maccani et al. 2014). In addition, CHO cell lines expressing well understood recombinant protein were mostly used, which might not be indicative when comparing miRNA profile dynamics in CHO cell lines secreting difficult-to-express (DTE) lysosomal enzymes. Studies
that have evaluated miRNA global expression changes in CHO cells secreting DTE protein (e.g. EPO-Fc fusion protein) either used microarray technology for analysis or steady-state cultivation of the CHO cell lines (Maccani et al. 2014; Klanert et al. 2016); microarray technology doesn’t have the high sensitivity of NGS for low (or rare) and high gene abundance (Hurd and Nelson 2009) just as steady-state cultivation may not give a true picture of miRNA expression changes as in a traditional fed-batch process. In this study, we sought to understand miRNA/mRNA expression changes in CHO cell lines (CHO-lys-low producer, CHO-lys-high producer, CHO-non-producer) secreting a DTE lysosomal protein. Furthermore, CHO-lys-low producer exhibited a 2 fold increase in growth rate compare to CHO-lys-high producer from the lag phase up to the exponential phase during cell culture cultivation (Figure 2A). In the production of therapeutic protein this phenomenon holds true because the increase in protein synthesis has been seen to be inversely proportional to cell growth (Loh et al. 2014). The lower growth rate of CHO-lys-high producer may be compensated with a higher productivity since most of the cells machineries are geared towards protein synthesis (Lee et al. 1998). The productivity difference between the low and high producer clones may be due to differences in their mRNA transcript levels, miRNA expression differences or differences in the site of integration of the transgene. We conducted a qRT-PCR experiment to ascertain the mRNA levels of the lysosomal protein secreted by these cell lines and found that there was no significant differences between the two clones at the transcript level (Figure 7B). In addition, clonal variation or CHO cell genetic heterogeneity could also a play a role in the differences between the productivity yields of these clones (Scarcelli et al. 2018; Derouazi et al. 2006). In order to develop a meaningful and reproducible results, these cell lines were grown in a 5L Dasgip bioreactor and cultivated for 14 days under a fed-batch biphasic (temperature shift on day 5) condition and samples were taken on day 3, 5, 7 and 10 for RNA-seq and small RNA-seq (miRNA)
processing. To understand the global mRNA and miRNA profile differences between CHO-lys-low and CHO-lys-high producers, we identified 500 DE mRNA (FDR<0.05 & FC>2 or <0.5) and 35 DE miRNA (FDR<0.05).

Among the 35 DE miRNAs, cgr-miR-31-5p was found to be 131 fold upregulated in CHO-lys-high producer compared to CHO-lys-low producer by taking the mean fold change at four different time points (day 3, 5, 7, & 10) during cultivation. The high expression profile of miR-31-5p in CHO-lys-high producer could explain its high pro-productive nature and/or high volumetric yield relative to the CHO-lys-low producer. Joon-Ho Cho et al. were able to show that the upregulation of miR-31 by histone deacetylase inhibitor (HDACi) caused an induction of cellular senescence in cancer cells (Cho, Dimri, and Dimri 2015). This could explain the 2 fold difference in growth rate between CHO-lys-high and CHO-lys-low producers. However, from the growth correlating result, miR-31 was positively correlated with growth (but not unique to only CHO-lys-high, also found in CHO-non-producer but not in CHO-lys-low) and negatively correlated with titer (Table 9 & 10). This could indicate the pleiotropic nature of miRNAs expression in different environmental conditions (Fischer, Handrick, et al. 2015).

Looking at the DE miRNAs associated with growth (data not shown), we observed that there were more DE miRNAs in the CHO-non-producer and CHO-lys-producer than in CHO-high producer. This observation is in contrast with the result by Maccani et al (2014) in which they observed much increase in mature miRNAs expression profiles in producer cell lines (Maccani et al. 2014). Interestingly, most of the upregulated miRNAs (e.g. miR-409-3p, let-7f) have been shown to cause a reduction in cellular proliferation when overexpressed. This is could be indicative why CHO-lys-high producer exhibited a growth rate that is 2 fold lesser compared to CHO-lys-low. Taken together, we have been able to profile the mRNA and
miRNA expression in CHO cell lines that differ in their productivity and in addition identified potential miRNAs as targets for CHO cell engineering.

{BIBLIOGRAPHY}

DECLARATIONS

Availability of data and material: The datasets supporting the conclusions of this article are included within the article.

Funding: This study was funded by BioMarin Pharmaceutical Inc.

Authors’ contributions: All authors participated in the conceptualization of the study and contributed to the experimental design and data analysis. MA performed the investigation and wrote the manuscript and all authors contributed to its edition.


Acknowledgments: The authors would like to thank JL, VA, and TC at BioMarin Pharmaceutical Inc. for their valuable suggestions and support throughout the execution of the study.

REFERENCES.


Baek, Daehyun, Judit Villén, Chanseok Shin, Fernando D. Camargo, Steven P. Gygi, and


Dyson, Michael R. 2016. “Fundamentals of Expression in Mammalian Cells.” In Advanced


Galatenko, Vladimir V., Alexey V. Galatenko, Timur R. Samatov, Andrey A. Turchinovich, Maxim Yu Shkurnikov, Julia A. Makarova, and Alexander G. Tonevitsky. 2018. “Comprehensive Network of miRNA-Induced Intergenic Interactions and a Biological


Inwood, Sarah, Eugen Buehler, Michael Betenbaugh, Madhu Lal, and Joseph Shiloach. 2017. “Identifying HIPK1 as Target of miR-22-3p Enhancing Recombinant Protein
Production From HEK 293 Cell by Using Microarray and HTP siRNA Screen.”


Johari, Yusuf B., Scott D. Estes, Christina S. Alves, Marty S. Sinacore, and David C.


**Table 7**: Parameters used for next-generation sequencing

<table>
<thead>
<tr>
<th></th>
<th>RNA-deq</th>
<th>small RNA-seq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample type</td>
<td>suspension cells</td>
<td>suspension cells</td>
</tr>
<tr>
<td>library</td>
<td>48</td>
<td>48</td>
</tr>
<tr>
<td>Profiling Platform</td>
<td>illumina HiSeq 4000</td>
<td>illumina HiSeq 4000</td>
</tr>
<tr>
<td>Read length</td>
<td>Paired-end 100 bp</td>
<td>Single-end 50 bp</td>
</tr>
<tr>
<td>Data output</td>
<td>30 million reads per sample</td>
<td>10 million reads per sample</td>
</tr>
</tbody>
</table>

**Table 8**: Relevant bioprocess characteristics of CHO cell lines used for cell culture in 5L bioreactor

<table>
<thead>
<tr>
<th></th>
<th>CHO-lys-low</th>
<th>CHO-lys-high</th>
<th>CHO-non-producer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth</td>
<td>2x</td>
<td>X</td>
<td>2x</td>
</tr>
<tr>
<td>Specific productivity</td>
<td>X</td>
<td>3.0x</td>
<td></td>
</tr>
<tr>
<td>Titer</td>
<td>x</td>
<td>1.8x</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>---</td>
<td>-----</td>
<td></td>
</tr>
</tbody>
</table>

**Table 9:** Growth correlating DE miRNAs unique to CHO-lys-high producer

<table>
<thead>
<tr>
<th>Positively correlation</th>
<th>Negative correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>cgr-let-7d-3p</td>
<td>novel-mir-5393</td>
</tr>
<tr>
<td>cgr-miR-15b-5p</td>
<td>cgr-miR-486-5p</td>
</tr>
<tr>
<td></td>
<td>cgr-miR-212</td>
</tr>
<tr>
<td></td>
<td>novel-mir-12388</td>
</tr>
<tr>
<td></td>
<td>novel-mir-10785</td>
</tr>
<tr>
<td></td>
<td>novel-mir-13139</td>
</tr>
</tbody>
</table>

**Table 10:** Titer correlating DE miRNAs unique to CHO-lys-high producer

<table>
<thead>
<tr>
<th>Positively correlation</th>
<th>Negative correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>cgr-miR-10b-5p</td>
<td>cgr-miR-31-5p</td>
</tr>
<tr>
<td>cgr-miR-146b-5p</td>
<td>cgr-let-7d-3p</td>
</tr>
<tr>
<td>cgr-miR-409-3p</td>
<td>cgr-miR-181b-5p</td>
</tr>
<tr>
<td>cgr-miR-32-5p</td>
<td>cgr-miR-17-3p</td>
</tr>
<tr>
<td>cgr-let-7f</td>
<td>novel-mir-2923</td>
</tr>
</tbody>
</table>

**Table 11:** Gene ontology and Functional annotation of miR-31-5p target genes predicted by at least 3 in silico tools in combination with DE genes from the RNA-Seq data along with potential genes for downstream analyses
<table>
<thead>
<tr>
<th>Functional group</th>
<th>FDR (&lt;0.2)</th>
<th># Genes</th>
<th>Relevant genes to bioprocess</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytoplasm</td>
<td>0.0002</td>
<td>151</td>
<td>MAPK14, NECAB3, MAP3K2, PPP2R5A, SMAD3, UBE2I3, EIF5A2</td>
</tr>
<tr>
<td>Membrane</td>
<td>0.046</td>
<td>145</td>
<td>TMED10, PPP2R5A, EIF5A2</td>
</tr>
<tr>
<td>Cytosol</td>
<td>0.054</td>
<td>48</td>
<td>MAPK14, MAP3K2, PPP2R2A, PPP2R5A</td>
</tr>
<tr>
<td>Endoplasmic reticulum memb.</td>
<td>0.065</td>
<td>25</td>
<td>TMED10, NECAB3, EIF5A2</td>
</tr>
<tr>
<td>Protein phosphatase type 2A regulator activity</td>
<td>0.076</td>
<td>5</td>
<td>PPP2R2A, PPP2R5A</td>
</tr>
<tr>
<td>Actin binding</td>
<td>0.084</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Phosphatase activity</td>
<td>0.094</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Myosin complex</td>
<td>0.094</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Protein phosphatase regulator activity</td>
<td>0.1</td>
<td>5</td>
<td>PPP2R2A, PPP2R5A</td>
</tr>
<tr>
<td>ATP binding</td>
<td>0.11</td>
<td>41</td>
<td>MAPK14, MAP3K2, UBE2I3</td>
</tr>
<tr>
<td>Protein binding</td>
<td>0.12</td>
<td>88</td>
<td>MAPK14, NECAB3, PPP2R5A, SMAD3</td>
</tr>
<tr>
<td>Endoplasmic reticulum</td>
<td>0.16</td>
<td>36</td>
<td>TMED10, NECAB3, EIF5A2</td>
</tr>
</tbody>
</table>

**Table 12:** Genes selected for downstream analysis from the functional annotation result of miR-31 target genes derived from prediction by at least 3 in silico tools and DE genes from RNA-Seq data.
<table>
<thead>
<tr>
<th>Genes selected</th>
<th># Times represented</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAPK14</td>
<td>4</td>
</tr>
<tr>
<td>TMED10</td>
<td>3</td>
</tr>
<tr>
<td>NECABS</td>
<td>4</td>
</tr>
<tr>
<td>MAP3K2</td>
<td>3</td>
</tr>
<tr>
<td>PPP2R2A</td>
<td>3</td>
</tr>
<tr>
<td>PPP2R5A</td>
<td>6</td>
</tr>
<tr>
<td>SMAD3</td>
<td>2</td>
</tr>
<tr>
<td>UBE2F</td>
<td>2</td>
</tr>
<tr>
<td>EIF5A2</td>
<td>4</td>
</tr>
</tbody>
</table>

**Table 13:** Functions of potential miR-31-5p target genes with relevance to bioprocess phenotypes
<table>
<thead>
<tr>
<th>Genes</th>
<th># Times represented gene</th>
<th>Functions</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAPK14</td>
<td>4</td>
<td>augmented expression confer repression of lung cancer growth in vitro and in vivo. Negatively regulate cell growth and division. Targeting MAPK7 with miR-143 has been shown to enhance protein production</td>
<td>Qi and Elion 2005; Raman et al. 2007; Keshet and Seger 2010; Schoellhorn et al, 2017</td>
</tr>
<tr>
<td>TMED10</td>
<td>3</td>
<td>Involved in vesicular protein trafficking. Mainly functions in the early secretory pathway</td>
<td>Denzel et al, 2000</td>
</tr>
<tr>
<td>NECAB3</td>
<td>4</td>
<td>Promotes activation of hypoxia inducible factor-1 (HIF-1) to use glycolysis even at normal oxygen levels, or normoxia.</td>
<td>Nakaoka et al, 2016</td>
</tr>
<tr>
<td>MAP3K2</td>
<td>3</td>
<td>Regulate processes such as cell proliferation, cell differentiation, and cell death in eukaryotes from yeast to humans</td>
<td></td>
</tr>
<tr>
<td>PPP2R2A</td>
<td>3</td>
<td>Augmented expression confer repression of lung cancer growth in vitro and in vivo. Negatively regulate cell growth and division</td>
<td>Xi Liu et al, 2010</td>
</tr>
<tr>
<td>PPP2R5A</td>
<td>6</td>
<td>PP2A phosphatase complex, a key negative regulator of the MAP kinase pathway. PP2A is regulated by the kinase mTOR. Both enzymes affect phosphorylation status of ribosomal protein S6</td>
<td>Mazur et al, 2014</td>
</tr>
<tr>
<td>SMAD3</td>
<td>2</td>
<td>The Smad family proteins are critical components of the TGF-β signaling pathways. TGF-β are inhibitory growth factors and regulates transcriptional responses</td>
<td>Kaji et al, 2001</td>
</tr>
<tr>
<td>UBE2F</td>
<td>2</td>
<td>Ubiquitination was initially identified as the main process for protein degradation. miR-30 overexpression showed decreased Skp2 &amp; Ube2j mRNA levels which enhanced protein production in CHO cells</td>
<td>Nandi et al, 2006; Fischer et al, 2015</td>
</tr>
<tr>
<td>EIF5A2</td>
<td>4</td>
<td>suppression of EIF5A2 in HGC27 cells led to significant decreases in cell proliferation. EIF5A2 positively regulated cyclin D1 and cyclin D3, which play important roles in cell proliferation and cell cycle regulation</td>
<td>Meng et al, 2015</td>
</tr>
</tbody>
</table>
**Figure 2:** Growth profiles of CHO-lys-low producer, CHO-lys-high producer and CHO-non-producer (parental cell line) grown in 5L bioreactors in 3L w/v cell culture for 14 days (A) specific growth rate (B) Viable cell concentration (C). Viability (D) Integral viable cell density (IVCD)
Figure 3: Productivity profiles and offline Glucose concentration/pH trends of CHO-lys-low producer, CHO-lys-high producer and CHO-non-producer (parental cell line) grown in 5L bioreactors in 3L w/v cell culture for 14 days (A) Specific productivity (B) Titer (C) Glucose concentration (D) Offline pH
**Relative endogenous novel-miR-2471 expression @ D7**

**Relative endogenous miR-196b expression @ D7**

**Relative endogenous miR-15b expression @ D7**

**Relative endogenous miR-100 expression @ D7**

**Relative endogenous miR-409 expression @ D7**

**Relative endogenous miR-221 expression @ D7**

**Relative endogenous miR-486 expression @ D7**

**Relative endogenous miR-10a expression @ D7**

**Relative endogenous miR-31 expression @ D7**
Figure 4: qRT-PCR confirmation of 14 DE (FDR<0.05 & FC>2 or <0.5) miRNAs selected from 35 DE (FDR<0.5) miRNAs identified by small RNA-Seq over cell culture duration. MiRNA expression was normalized to U6 snoRNA. Error bars represent the standard error of mean (SEM) of three independent experiments (n=3). For statistical analysis, one-way ANOVA comparing the mean values of the mimics (* P < 0.05; ** P < 0.01; *** P < 0.001).
Figure 5: Quantitative real-time PCR (qRT-PCR) analyses. Cells were transiently transfected with 50 nM of miR-31-5p mimics or inhibitors in CHO-lys-low producer expressing DTE lysosomal enzyme. Cell lysates harvested 48 h post-transfection (A) Endogenous miR-31-5p levels post-transfection with miR-31-5p inhibitor (B) Endogenous miR-31-5p levels post-transfection of miR-31-5p mimics
**Figure 6:** Growth profile of CHO-lys-low producer post-transfection with (A) mouse miR-31-5p mimic (B) mouse miR-31-5p inhibitor (C) Product concentration (titer) of lysosomal enzyme secreted by CHO-lys-low producer post-transfection of mouse miR-31-5p inhibitor (D) Product concentration (titer) of lysosomal enzyme secreted by CHO-lys-low producer post-transfection of mouse miR-31-5p mimic (E) Product concentration (titer) of lysosomal enzyme secreted by CHO-lys-high producer post-transfection of mouse miR-31-5p mimic and inhibitor. Error bars represent the standard error of mean (SEM) of three independent experiments (n=3). Error bars represent the standard error of mean (SEM) of three independent experiments (n=3). For statistical analysis, student t-test was used (* P < 0.05; ** P < 0.01; *** P < 0.001).
Figure 7: (A) Functional annotation clustering of DE target genes of miR-31-5p predicted by at least 3 in silico tools in combination with miRNA-target gene result from the NGS data analyses. (B) Quantitative real-time PCR (qRT-PCR) analyses of day 3 RNA samples used for NGS for both CHO-lys-low and CHO-lys-high producer cell lines. mRNA levels of GALC.
Table 14: Primers used for miRNA and mRNA expression analyses using qRT-PCR

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Forward primer (5’ – 3’)</th>
<th>Reverse Primer (5’ – 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAPDH</td>
<td>AGCCACATCGCTCAGACAC</td>
<td>GCCCAATACGACAAATCC</td>
</tr>
<tr>
<td>U6 snoRNA</td>
<td>CTCGCTTCGGCAGCACA</td>
<td>AACGCTTCAAGAATTTGCGT</td>
</tr>
<tr>
<td>Lysosomal gene</td>
<td>proprietary</td>
<td>proprietary</td>
</tr>
<tr>
<td>cgr-miR-31-5p</td>
<td>AGGCAAGAUGCUUGCAGAUGCU</td>
<td>Qiagen universal reverse primer</td>
</tr>
</tbody>
</table>

Table 15: miRTarbase (v 7.0) list of validated targets of positively titer-correlating miRNAs unique to CHO-lys-high producer cell line

<table>
<thead>
<tr>
<th>miRNA</th>
<th>Validated targets</th>
<th>Gene description</th>
<th>Gene Functions</th>
<th>Validated method (miRTarbase)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>mmu-miR-10b-5p</td>
<td>Cdkn1a</td>
<td>cyclin-dependent kinase inhibitor 1A</td>
<td>Inhibits cellular proliferation mediated by p53/TP53 and cyclin dependent kinase activity preventing cell cycle progression</td>
<td>qPCR</td>
<td>Harper et al, 1993</td>
</tr>
<tr>
<td></td>
<td>Klf4</td>
<td>Kruppel-like factor 4</td>
<td>Inhibition increased cancer cell migration</td>
<td>qPCR</td>
<td>Ma Li, 2010</td>
</tr>
<tr>
<td></td>
<td>Hoxd10</td>
<td>Homeobox D10</td>
<td>Sequence-specific TF. Inhibition represses metastasis in mammary tumor</td>
<td>WB</td>
<td>Ma et al, 2010</td>
</tr>
<tr>
<td>mmu-miR-146b-5p</td>
<td>Sirt1</td>
<td>Sirtuin 1</td>
<td>Maintains metabolic homeostasis in adipose tissues</td>
<td>qPCR, WB, Reporter Assay</td>
<td>Ahn et al, 2013</td>
</tr>
<tr>
<td></td>
<td>Nfb1</td>
<td>Nuclear factor of kappa light polypeptide gene enhancer in B-cells</td>
<td>decreased expression of siah2 (miR-146b target), inhibits ubiquination of TRAF proteins upregulatin NF-kB</td>
<td>WB</td>
<td>Nata et al, 2013</td>
</tr>
<tr>
<td></td>
<td>Notch1</td>
<td>Notch 1</td>
<td>Knockdown promoted stem cell proliferation in mouse</td>
<td>Reporter assay</td>
<td>Xiao et al, 2015</td>
</tr>
<tr>
<td></td>
<td>Smad4</td>
<td>SMAD family member 4</td>
<td>Are critical components of the TGF-β signaling pathways. TGF-β are inhibitory growth factors and regulates</td>
<td>Reporter assay, qPCR</td>
<td>Kaji et al, 2001</td>
</tr>
<tr>
<td>Genotype</td>
<td>Description</td>
<td>Function</td>
<td>Reporter Assay</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
<td>----------</td>
<td>----------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>IRF5</td>
<td>interferon regulatory factor 5</td>
<td>Key transcription factor for macrophage differentiation</td>
<td>Reporter assay</td>
<td>Peng et al, 2016</td>
<td></td>
</tr>
<tr>
<td>mmu-miR-409-3p</td>
<td>nil</td>
<td>Suppresses cellular proliferation (Li et al, 2012)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mmu-let-7f</td>
<td>nil</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mmu-miR-192</td>
<td>H3f3b H3 histone, family 3B</td>
<td>nucleosome structural function</td>
<td>qPCR, Reporter Assay</td>
<td>Lu et al, 2017</td>
<td></td>
</tr>
<tr>
<td>mmu-miR-342-3p</td>
<td>nil</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mmu-miR-183</td>
<td>Taok1 TAO kinase 1</td>
<td>Activation of MAPK pathway in response to stress</td>
<td>qPCR</td>
<td>Patel et al, 2013</td>
<td></td>
</tr>
<tr>
<td>Zeb2</td>
<td>zinc finger E-box binding homeobox 2</td>
<td>encodes repressors of E-cadherin</td>
<td>Reporter assay</td>
<td>Oba et al, 2013</td>
<td></td>
</tr>
<tr>
<td>Lrp6</td>
<td>low density lipoprotein receptor-related protein 6</td>
<td>impairs Wnt/β-catenin signalling pathway</td>
<td>qPCR, WB, Reporter Assay</td>
<td>Chen et al, 2014</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 5

Inhibitory impact of miR-23a/miR-377 on SUMF1 in CHO cells enhances difficult-to-express recombinant lysosomal sulfatase activity

Ifeanyi Michael Amadi¹², Vishal Agrawal¹, Terri Christianson¹, Cameron Bardliving², Parviz Shamlou², Jonathan Lebowitz¹

¹ BioMarin Pharmaceutical Inc. 105 Digital Dr. Novato, California 94949, USA
² Keck Graduate Institute. 535 Watson Dr. Claremont, California 91711, USA

ABSTRACT

Difficult-to-express (DTE) recombinant proteins like multi-specific proteins, DTE monoclonal antibodies and lysosomal enzymes, have seen difficulties in manufacturability using Chinese hamster ovary (CHO) cells and other mammalian cells as production platforms. CHO cells are preferably used for protein production because of their innate ability to secrete human-like recombinant proteins with post-translational modification, resistance to viral infection and familiarity with drug regulators. However, despite huge progress made in engineering CHO cells for high volumetric productivity, DTE proteins like recombinant lysosomal sulfatase represent one of the poorly understood proteins. Furthermore, there are growing interest in the use of microRNA (miRNA) to engineer CHO cells expressing
DTE proteins to improve cell performance of relevant bioprocess phenotypes. To our knowledge, no research has been done to improve CHO cell production of DTE recombinant lysosomal sulfatase using miRNA. We identified miR-23a and miR-377 as miRNAs targeting SUMF1 using in silico prediction tools as SUMF1 is an activator of sulfatases. Transient inhibition of endogenous miR-23a/miR-377 significantly enhanced recombinant sulfatase enzyme specific activity in CHO cell without affecting cell growth. Though, inhibition of miR-23a/miR-377 had no significant effect on the mRNA and protein levels of SUMF1, overexpression of miR-23a/377 significantly reduced both the mRNA and protein levels of SUMF1. In summary, our data demonstrates the importance of using miRNA to optimize CHO cell line secreting DTE recombinant lysosomal sulfatase.

**Keywords**: Chinese hamster ovary (CHO cells), cell engineering, microRNA, lysosomal protein, sulfatase, miR-23a, miR-377

**Abbreviations**: CHO, Chinese hamster ovary; DTE, difficult-to-express; SUMF1, sulfatase modifying factor 1; miRNA, microRNA; UTR, untranslated region; FGE, formylglycine (FGly)-generating enzyme; qRT-PCR, real-time quantitative reverse-transcription PCR.
5.0 Introduction

The use of mammalian cells, most especially CHO, as a production platform for recombinant therapeutic proteins has increased for decades because of its ability to secrete proteins with similar post-translational modifications (PTM) like humans (Fliedl, Grillari, and Grillari-Voglauer 2015). Though significant improvements have been made at increasing specific and volumetric productivities of cells using media optimizations (Clincke et al. 2011; Buchsteiner et al. 2018), process improvements (Huang et al. 2010; Gagnon et al. 2011) and genetic engineering (Jazayeri et al. 2018), there have been limitations especially for difficult-to-express (DTE) proteins with very low production titers or improper folding compared to biomanufacturing standards. The production of recombinant therapeutic lysosomal proteins have not experienced the tremendous improvements in protein production as seen in most monoclonal antibodies, secreted alkaline phosphatase (SEAP), and other easy-to-express proteins. Though classical cell engineering of cells has been done by the overexpression or knockdown of protein coding genes, this approach introduces additional translational burden and cellular limitations.

In recent years, there has been growing interest in the use of microRNA (miRNA) to engineer CHO cells for improvement in cell performance of relevant bioprocess phenotypes (Jadhav et al. 2012; Niall Barron 2012; Jadhav et al. 2013; Stiefel et al. 2016). MiRNAs are short non-coding RNA molecules, about 18-24 nucleotide long; they negatively regulate gene expression post-transcriptionally via translation repression or mRNA degradation without adding any translational burden to the cell machinery unlike protein-coding genes. In addition, miRNA can regulate an entire network of genes (Galatenko et al. 2018) and it regulate gene expression by acting as a rheostat (Baek et al. 2008), partially reducing protein expression unlike small-interfering RNAs (si-RNAs) that most often completely obliterate their protein target. MiRNA regulate gene expression by binding to the 3’
untranslated region (UTR) of their mRNA target by using their “seed” region located at the 5’ region of the miRNA. Based on these desirable qualities of miRNA, scientists have engineered CHO cells to optimize relevant bioprocess phenotypes e.g., apoptosis (Druz et al. 2013), cell proliferation (Hackl et al. 2014), and specific productivity (Strotbek et al. 2013; Loh et al. 2014; Meyer et al. 2017; Fischer, Paul, et al. 2015; Loh, Yang, and Lam 2017; Inwood et al. 2017; Emmerling et al. 2016). However, most of the CHO cell lines used only secreted easy-to-express recombinant proteins and only of recent has miRNA engineering of CHO cell line expressing DTE proteins been tested. MiR-143 and miR-557 has been shown to enhance protein production and improve cell line development respectively in CHO cell lines expressing DTE monoclonal antibodies (Schoellhorn et al. 2017; Fischer et al. 2017; Strotbek et al. 2013). MiR-143 exhibited this attribute by repressing mitogen-activated protein kinase (MAPK7) and MAPK pathways has been shown to regulate gene transcription, protein synthesis and differentiation (Kyriakis and Avruch 2001). In addition, target gene for miR-557 has not been identified till date.

Previous study (Frankel et al. 2014) demonstrated that by inhibiting miR-95, which regulates the expression of residual sulfatase modifying factor 1 (SUMF1), sulfatase activity could be increased. In this study, Frankel et al. transfected human cancer lines and fibroblasts to demonstrate the effect of miR-95 which is a non-conserved (not found in mice) miRNA. SUMF1 plays a crucial role in the production of lysosomal sulfatase proteins, since it encodes Cα-formylglycine (FGly)-generating enzyme (FGE) which activates sulfatases by converting their cysteine residue to active FGly. In addition, inhibition of miR-23 has been shown to increase CHO cell productivity by increasing oxidative metabolism (Kelly et al. 2015). However, the CHO cell secretes an easy-to-express SEAP protein which may be unable to stretch the full capability of the miRNA engineering function.
In this study, we sought to identify and functionally test miRNA(s) that target SUMF1 and improve recombinant sulfatase specific activity. MiR-23a and miR-377 were identified to target SUMF1 using in silico computational predictive tools e.g., TargetScan, miRanda-mirSVR, Diana-microT-CDS, Diana-Tarbase and miRecords. Following this, we used an in-house CHO cell line stably expressing a recombinant lysosomal sulfatase enzyme for cell line engineering. Furthermore, transient functional inhibition of endogenous miR-23a/miR-377 using miR-23a/miR-377 antimiRs enhanced recombinant sulfatase enzyme specific activity in CHO-sulfatase cell without affecting cell growth. Though, inhibition of miR-23a/miR-377 had no significant effect on the mRNA and protein levels of SUMF1, overexpression of miR-23a/miR-377 significantly reduced both the mRNA and protein levels of SUMF1.

5.1 Materials and Methods

5.1.1 Cell culture

CHO suspension cells stably expressing sulfatase protein (CHO-sulfatase) were cultivated in CD CHO medium (Thermo Scientific, Waltham, MA, USA) and routinely propagated in 6-well plates (Life Technologies, Grand Island, New York) and maintained at 37°C, 5% CO2 and 80% humidity in a static incubator (Thermo Scientific, Waltham, MA, USA). Inoculation cells were sourced from cells maintained in 30-60 mL shake flask (Corning, Oneonta, NY, USA) at 37°C, 5% CO2 and 80% humidity in an orbital shaker incubator (Infors HT, Annapolis Junction, Maryland, USA) with agitation at 125 rpm. Cells were passaged every 3-4 days at a seeding density of 0.5 x 10^6 cells/mL. Cell concentration and viability were measured using the ViCell™ (Beckman Coulter, Brea, CA, USA) by the trypan blue exclusion method. HEK293, HT1080 and MCF7 cells were cultivated in DMEM medium supplemented with 10% FBS (VWR, Radnor, PA, USA), and 2 mM GlutaMAX™ (Thermo
Scientific, Waltham, MA, USA). The mammalian cells were routinely propagated in T75 flasks (Corning, Oneonta, NY, USA).

5.1.2 Transfection

Transient transfection of miRNA mimics, inhibitors and siRNAs (Qiagen, Germantown, MD, USA; Dharmacon, GE Healthcare, Lafayette, CO, USA) were carried out using Lipofectamine™ RNAiMax (Thermo Scientific, Waltham, MA, USA) transfection reagent and Opti-MEM I (Thermo Scientific, Waltham, MA, USA) reduced serum medium for complex formation at a final concentration of 50 nM, unless indicated differently. Anti-SUMF1 SMARTpool™ siRNAs (mouse si-SUMF1) (Dharmacon, GE Healthcare, Lafayette, CO, USA), which comprises of a mixture of 4 siRNAs, was used as a functional control and the scramble universal negative control (Sigma-Aldrich, Louis, MO, USA) as a non-targeting control (NT-siRNA).

5.1.3 RNA Isolation

Total RNA, including small RNA, was extracted using Quick-RNA Microprep kit (Zymo Research, Irvine, CA, USA) according to the protocol of the manufacturer from samples harvested from cell culture 48 days post-transfection. The concentration and purity of the RNA was determined by UV-spectrometry using a DS-11 spectrophotometer (DeNovix, Wilmington, DE, USA) by measuring absorbance at 230, 260, and 280 nm.
5.1.4 Real-time Quantitative Reverse-Transcription PCR (qRT-PCR)

Briefly, cDNA was synthesized from 1 μg or 500 ng (where applicable) of total RNA using miScript II kit (Qiagen, Germantown, MD, USA) for quantification of mature miRNA(s) expression and Superscript Vilo kit (Invitrogen, Thermo Scientific, Carlsbad, CA, USA) for gene expression quantification according to the protocol of the manufacturers. For the quantification of matured miRNAs and gene expression, qRT-PCR was carried out with 10⁻¹ diluted cDNA on a LightCycler 480 (Roche Molecular Diagnostics, Pleasanton, CA, USA) using Ssofast Evagreen® Supermix (BioRad, Hercules, CA, USA) according to the protocols of the manufacturer. GAPDH and U6 snoRNA were used for the normalization of gene expression and mature miRNA expression levels respectively. miScript Universal Reverse Primer (Qiagen, Germantown, MD, USA) was used for mature miRNA analysis. Details of primers used for qRT-PCR (supporting information Table S1). For the quantification of relative gene expression, the comparative C(T) method was employed.

5.1.5 Western blot analysis

Protein extraction on harvested cells was done using the M-PER mammalian Protein Extraction Reagent Kit (Thermo Scientific, Carlsbad, CA, USA) according to the protocol of the manufacturer. Halt Protease Inhibitor Cocktail (Thermo Scientific, Carlsbad, CA, USA) was mixed with the extraction solution at a ratio of 1:100 respectively. The total protein of the protein extract was determined using the bicinchoninic acid (BCA) protein assay (Thermo Scientific, Carlsbad, CA, USA) according to the protocol of the manufacturer. Cellular proteins were separated in a 4-12% SDS-polyacrylamide Nupage premade gels (Thermo Scientific) using a 1X MOPS SDS running buffer. Proteins were transferred from the gel to a nitrocellulose membrane using the iBlot 2 dry blotting system (Life Technologies).
according to the protocols of the manufacturer. The membranes were allowed to wash in a blocking buffer (Thermo Scientific) for 1 hour on a shaker at room temperature. The following primary antibodies: human SUMF1 (1:1000, TA337720, Origene, Rockville, MD, USA), mouse/human SUMF1 (1:500, MAB2779, R&D Systems), GAPDH (1:5000, PA1-988, Thermo Scientific), were diluted in blocking buffer and used to probe the membrane overnight at 4\(^{0}\) C on a shaker. Thereafter, the membranes were probed with fluorochrome-conjugated secondary antibodies (LICOR Biosciences, Lincoln, Nebraska, USA) and detected with Odyssey CLx Imaging System (LICOR Biosciences, Lincoln, Nebraska, USA) after 1 hour of incubation.

### 5.1.6 Sulfatase enzymatic activity and Elisa assays

For determination of enzymatic activity of sulfatase in sample supernatant, 4-methylumbelliferyl (4-MU) sulfate was used as substrate and incubated in 0.2M NaOAc buffer (pH 5.6). 0.4M phosphate buffer (pH 6.7) was used to stop reaction and fluorescence was read at 355 nm (excitation) and 460 nm (emission) on a SpectraMax M3 plate reader (Molecular Devices, San Jose, CA, USA). Sulfatase protein concentration of sample supernatant was determined using an Elisa assay. Sulfatase was captured in a binding 96-well plate, washed and blocked with a washing and blocking buffer respectively. Anti-mouse HRP antibody was used to bind detection antibody and reaction detected using QuantaBlu fluorogenic peroxidase substrate kit (Thermo Scientific, Carlsbad, CA, USA) according to the manufacturer’s protocol. Fluorescence was read at 325 nm (excitation) and 420 nm (emission) on a SpectraMax M3 plate reader (Molecular Devices, San Jose, CA, USA)
5.1.7 Bioinformatics

MiRNAs targeting SUMF1 mRNA were determined using computational prediction tools. The following tools were employed: TargetScan (v.7.1), miRanda-mirSVR (Aug’10), Diana-microT-CDS (v5.0), Diana-Tarbase (v.8) in combination with miRecords (April 27, 2013). MiRNAs predicted to target SUMF1 in mouse was used since there are no current target prediction tool for CHO cells. Only potential miRNAs predicted to target SUMF1 mRNA by at least four prediction tools were considered for downstream functional analysis.

5.2 Results

5.2.1 In silico identification of putative miRNAs targeting SUMF1

Frankel et al (2014) showed that non-conserved miR-95 have an indirect effect on sulfate metabolism by regulating SUMF1 gene (Frankel et al. 2014). In addition, they found out that by inhibiting miR-95, the enzymatic activity of Arylsulfatase B (ARSB) is increased. To improve the specific activity of an in-house CHO cell expressing a sulfatase, we set out to identify conserved miRNA(s) targeting SUMF1. Overall, 62 microRNAs (Table 17) were predicted altogether by the four bioinformatics tools used and 234 (includes the 62 microRNAs) by miRecords. Only 6 of the 62 microRNAs predicted by all four prediction tools were selected for downstream functional studies. Two of these miRNAs (miR-23a-3p and miR-377-3p) predicted to target SUMF1 were pursued further.

5.2.2 Characterization of different mammalian cells to determine appropriate SUMF1 antibody

In order to determine the appropriate SUMF1 antibody for this study we went ahead to test different mammalian cell lines for the expression of endogenous SUMF1 protein. A total of
four different mammalian cell types were used: CHO, HT1080, HEK 293, and MCF7. For the CHO cell lines, three different cell types were used. One of the CHO cell line clone stably expresses a recombinant sulfatase and human SUMF1, the other stably expresses only a recombinant sulfatase and the last is a parental CHO cell line expressing no recombinant protein. In addition, the three different SUMF1 antibodies were tested on the mammalian cell lines in both reduced and non-reduced condition in a Western blot. As expected, the human anti-SUMF1 antibody only detected the human recombinant SUMF1 stably expressed by one of the CHO cell lines in both reduced and non-reduced conditions (Figure 8A). In like manner, mouse/human anti-SUMF1 antibody only detected the human recombinant SUMF1 stably expressed by one of the CHO cell lines in both reduced and non-reduced conditions except the detection of a small SUMF1 band in HEK 293 cells in reduced condition. However, mouse anti-SUMF1 antibody detected endogenous SUMF1 in all cell line tested except for HEK293 in a reduced condition. However, in non-reduced condition, only the HEK293 and MCF7 endogenous SUMF1 were not detected. In addition, si-hSUMF1 (siRNA) optimal concentration for transfection was also tested (Figure 14A-B).

5.2.3 Transient transfection of miR-23a and miR-377 mimics reduced SUMF1 protein and mRNA levels with no effect on sulfatase activity

We employed transient transfection to investigate the effect of the two miRNA mimics on sulfatase enzyme activity in a CHO cell line stably secreting lysosomal sulfatase. We observed that both mimics caused a reduction in SUMF1 protein (Figure 10A) and mRNA (Figure 10B) levels. The reduction in SUMF1 protein 48 h post miRNA transfection was more pronounced in cells transiently transfected with miR-23a compared to the scramble. Surprisingly, the mimics had no significant effect on the sulfatase enzyme titer (Figure 9B) and specific activity (Figure 9C). All cells transfected with miRNA mimic had similar cell growth profiles compared to the scramble (Figure 9A). Next, we tested the effect of various
concentrations of miR-23a mimics using the following concentrations: 10, 25, 50, 100 and 200 nM. Overall, miR-23a titration concentrations resulted in a decrease in protein (Figure 13B) and mRNA (Figure 13C) levels of SUMF1 with no effect on growth (Figure 13A). However, the reduction in mRNA levels from 10, 50 and 200 nM of miR-23a mimics was not significant and including 100 nM of miR-23a mimic at the protein level. Increased levels of mature miR-23a-3p and/or miR-377-3p were observed 48 h post-transfection following transient overexpression of miR-23a and/or miR-377 respectively (Figure 10C & Figure 13D).

5.2.4 Transient inhibition of miR-23a and miR-377 enhances recombinant sulfatase enzyme specific activity in CHO cell line

We sought to investigate the effect of transiently knocking down miR-23a and miR-377 expression on the specific activity of sulfatase stably secreted in a CHO cell line. To do this, mouse and CHO cells antimiRs were used knowing that they exhibited similar mature miR-23a and miR-377 sequences (Figure 8B). Surprisingly, 48 h post-transfection, there was no striking effect on the protein levels of SUMF1 except for CHO cells transfected with si-SUMF1 as a positive control in which SUMF1 levels were highly depleted (Figure 12A). Likewise, SUMF1 mRNA levels showed no significant differences with the scramble control except for cgr-miR-377 that showed a significant decrease in SUMF1 mRNA reduction compared to scramble control (Figure 12B). However, 72 h post-transfection all antimiRs tested showed a significant increase in sulfatase specific activity except for anti-mmu-miR-377-3p (Figure 11C) with no effect on growth (Figure 11A). In addition, there was no significant differences between the sulfatase titer of all antimiRs tested compared to scramble except for anti-cgr-miR-377-3p (Figure 11B). Furthermore, as a confirmation of functional and efficient transfection of antimiRs and siRNAs, mature miRNA levels analyzed
48 h post transfection showed downregulation of endogenous miR-23a and miR-377 compared to scramble (Figure 12C).

5.3 Discussion

Mammalian cells has become the traditional and classic cell factories used in the production of biopharmaceutical recombinant therapeutic proteins (Dyson 2016) because of their ability to make complex glycosylated proteins and familiarity with biotech regulatory bodies. Among mammalian cells, CHO cell lines most often represent the mammalian cell of choice for their ease of use and reduced susceptibility to viral infection (Berting, Farcet, and Kreil 2010; Dumont et al. 2016). However, due to growing interest in the demand and development of difficult-to-express (DTE) recombinant proteins such as bispecific antibodies, fusion proteins, and lysosomal proteins, current CHO cell lines may require optimization in order to operate at their optimum capacities. Though cell engineering approaches, such as single gene engineering (Onitsuka et al. 2018), has been employed to improve CHO cell performance there have been limitations associated with this (Le Fourn et al. 2014). Recently, there has been growing interest in the use of miRNAs to engineer CHO cells for relevant bioprocess phenotypes (Bratkovič et al. 2012), including applications in transient protein expression (Meyer et al. 2017) and cell line development (Fischer et al. 2017). MiRNAs are short non-coding RNAs that are 18-21 nucleotide long and each miRNA regulate gene expression by fine tuning the expression of more than one gene, thereby regulating more than one pathway concomitantly (Fischer, Handrick, et al. 2015). Although, much of the knowledge about miRNAs have stemmed from research on its regulatory effects on diseases (most especially cancer), research on its relevance in cell line engineering for the biomanufacturing of recombinant proteins have started gaining credence. MiRNA
engineering of CHO cells application has been demonstrated to impact relevant bioprocessing phenotypes like apoptosis (Druz et al. 2011), proliferation (N. Barron et al. 2011), cell specific productivity (Inwood et al. 2017) and metabolism (Kelly et al. 2015). Furthermore, current works are looking at using miRNA to engineer mammalian cells secreting DTE proteins for desirable and relevant bioprocess phenotypes (Xiao et al. 2015; Fischer et al. 2017; Schoellhorn et al. 2017). To our knowledge, this is the first study where miRNA has been used to engineer CHO cell line stably expressing a DTE recombinant therapeutic lysosomal protein – sulfatase. Lysosomal storage disorders (LSD) are genetic diseases caused by the accumulation of lysosomal substrates in the lysosome due to mutation in genes encoding lysosomal enzymes responsible for degrading accumulated substrates in cells (e.g., glycoaminoglycans –GAGs) (Platt 2018). An example of these diseases among others is multiple sulfatase deficiency (MSD) caused by mutations in sulfatase modifying factor 1 (SUMF1), a gene that encodes formylglycine –generating enzyme (FGE) (Garavelli et al. 2014). Interestingly, SUMF1 has been identified as a direct target of non-conserved miR-95 (Frankel et al. 2014). The authors discovered that by inhibiting miR-95, residual increase in SUMF1 protein levels in MSD fibroblast cells led to an increase in lysosomal sulfatase activity. SUMF1 plays a crucial role in activating sulfatases by converting cysteine residue in the consensus sequence of sulfatases to active formyl glycine (FGly) (Landgrebe et al. 2003). Although SUMF1 is conserved both in prokaryotes and eukaryotes (Landgrebe et al. 2003), miR-95 is only conserved in higher mammals but not in mice. In principle, since SUMF1 is conserved there might exist other conserved miRNA(s) in mice that might regulate its expression. We sought to identify miRNA(s) that regulate SUMF1 and the impact on recombinant therapeutic lysosomal sulfatase specific activity in CHO cell.
5.3.1 Transient inhibition of miR-23a and miR-377 increases lysosomal recombinant sulfatase specific activity in CHO cell

MiR-23a and miR-377 were identified as putative miRNAs regulating SUMF1 by using in silico computational prediction tools: TargetScan, miRanda-mirSVR, Diana-microT-CDS, Diana-Tarbase and miRecords. We sought to demonstrate functionally if these miRNAs might improve the specific activity of recombinant lysosomal sulfatase constitutively expressed in a CHO cell line. Transient inhibition of miR-23a and miR-377 increased the specific activity of recombinant lysosomal sulfatase protein by 1.2 – 1.3-fold increase compared to negative control scramble (Figure 11C) without affecting cell growth (Figure 11A). It seems 1.2 – 1.3-fold increase in sulfatase specific activity appears to be moderate compared to the scramble control, this could be because transient inhibitory effect of antimiRs dwindle over time. In addition, similar fold increase in sulfatase activity has been seen after overexpressing SUMF1 gene in a work by Fraldi et al. (2007). In this conjunction, a stable inhibition of miR-23a/miR-377 could further boost the specific activity of lysosomal sulfatase protein constitutively expressed in CHO cells. Notably, in a different study (Kelly et al. 2015), miR-23 was found to enhance the specific productivity of CHO cell expressing an easy-to-express human secreted alkaline phosphatase (SEAP) protein without affecting cell growth. The authors went further to demonstrate that the increased specific productivity of CHO cells after stably inhibiting miR-23 was as a result of enhanced oxidative phosphorylation through the TCA cycle. In an earlier study, mitochondrial glutaminase levels in human P-493B lymphoma cells and PC3 prostate cancer cells increased after downregulation of miR-23a (Gao et al. 2009). Glutamate, which is a product of glutamine breakdown in the presence of glutaminase, is funneled via the TCA cycle for ATP production which is very vital for protein synthesis. Furthermore, in a different study, overexpression of miR-377 caused a 60% reduction in MAP3K7 protein levels (Zehavi et al. 2015) and
regulation of MAPK7 gene in CHO by miR-143 has been shown to enhance the expression of a DTE protein (Schoellhorn et al. 2017). Hypothetically, after transient inhibition of miR-23a/miR-377 we expected an increase in SUMF1 mRNA and protein levels instead we observed no significant difference in SUMF1 protein (Figure 12A) and mRNA (Figure 12B) levels compared to the scramble control. This might be an indication that the right amount of SUMF1 as needed by the internal cellular mechanism is just right enough to increase the specific activity of recombinant lysosomal sulfatases while inhibiting the effect of endogenous miR-23a/miR-377 negatively regulating SUMF1 transcripts. Surprisingly, when miR-23a and miR-377 were transiently overexpressed, the SUMF1 protein (Figure 10A) and mRNA (Figure 10B) in CHO cells were significantly decreased compared to the scramble control. This phenomenon was also confirmed when we performed a titration of miR-23a mimics concentration and each concentration resulted in a decrease in mRNA (Figure 13C) and protein levels (Figure 13B) of SUMF1 with no effect on growth (Figure 13A). Though the reduction in mRNA using 10, 50 and 200 nM of miR-23a mimics were not significant (Figure 13C) and including 100 nM of miR-23a mimic for protein levels (Figure 13B). However, the reduction in mRNA and protein levels of SUMF1 as a result of miR-23a/miR-377 overexpression did not cause any decrease in recombinant lysosomal sulfatase specific activity. The explanation could be that there are other internal cellular mechanism that maintains a particular threshold of SUMF1 levels (Sardiello et al. 2005) and the optimal amount of SUMF1 required to activate sulfatases may vary across the different types of lysosomal sulfatases (Fraldi et al. 2007). In our case, the optimal SUMF1 required for activation of our sulfatase enzyme of interest might be low which might explain the reason there was no change in specific activity of the lysosomal sulfatase following the overexpression of miR23a/miR-377. Taken together, these data demonstrate the relevance
of miR-23a and miR-377 in cell engineering of CHO cells expressing DTE recombinant therapeutic lysosomal sulfatase protein.

{BIBLIOGRAPHY}

DECLARATIONS

Availability of data and material: The datasets supporting the conclusions of this article are included within the article.

Funding: This study was funded by BioMarin Pharmaceutical Inc.

Authors’ contributions: MA, TC, VA, JL participated in the conceptualization of the study and contributed to the experimental design and data analysis. MA performed the investigation and wrote the manuscript and all authors contributed to its edition.

Competing interests: MA, TC, VA, JL are employees and TC, VA, JL stockholders of BioMarin Pharmaceutical Inc.

Acknowledgments: The authors would like to thank JL, VA, and TC at BioMarin Pharmaceutical Inc. for their valuable suggestions and support throughout the execution of the study.

REFERENCES.


Inwood, Sarah, Eugen Buehler, Michael Betenbaugh, Madhu Lal, and Joseph Shiloach. 2017. “Identifying HIPK1 as Target of miR-22-3p Enhancing Recombinant Protein
Production From HEK 293 Cell by Using Microarray and HTP siRNA Screen.”


FIGURES

A

<table>
<thead>
<tr>
<th>ANTIBODIES TESTED</th>
<th>REDUCED CONDITION</th>
<th>NON-REDUCED CONDITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human anti-SUMF1</td>
<td>SUMF1</td>
<td>SUMF1</td>
</tr>
<tr>
<td>Mouse anti-SUMF1</td>
<td>SUMF1</td>
<td>SUMF1</td>
</tr>
<tr>
<td>Mouse/human anti-SUMF1</td>
<td>SUMF1</td>
<td>SUMF1</td>
</tr>
</tbody>
</table>

B

Conserved
Position 787-794 of SUMF1 3’ UTR 5’ ...CCCAAGCCUUCAAAGAAUGUGAA
mmu-miR-23a-3p 3’ CCUUUGGGACCCGUUACACUA

Conserved
Position 480-487 of SUMF1 3’ UTR 5’ ...GCUAUUUUUUGCCA - UGUUGUGAA
mmu-miR-377-3p 3’ UGUUUAACGGAAAACACACUA

Poorly conserved
Position 875-881 of SUMF1 3’ UTR 5’ ...AGACUCCAUCAGGGUGUGUGAU
mmu-miR-377-3p 3’ UGUUUAACGGAAAACACACUA

Figure 8: Characterization of different mammalian cells to determine appropriate SUMF1 antibody (A) Western blot detection of endogenous SUMF1 protein 72 h post transfection. In addition, SUMF1 protein detection in CHO-sulfatase cell stably expressing human SUMF1 was also tested (B) Predicted binding sites of miR-23a and miR-377 on SUMF1 3’ UTR. Bases in red represents the binding site of the miRNAs seed region to the 3’ UTR of SUMF1 (TargetScan.org).
Figure 9: Transient transfection of 50 nM miR-23a and miR-377 mimics in CHO suspension cells stably expressing a recombinant difficult-to-express (DTE) lysosomal sulfatase protein. Cells were cultured in a 6-well plate using a batch process and kept in a static incubator. (A) Viable cell density (VCD) and viability 72 h post transfection (B) Sulfatase titer and (C) specific activity determined by an Elisa and activity assay respectively. Samples were taken from the culture supernatants 72 h post-transfection and data represented as fold change normalized to scramble (NT-siRNA) control. Anti-SUMF1 siRNA (si-SUMF1) was used as a positive control. Error bars represent the standard error of mean (SEM) of three independent experiments (n=3). For statistical analysis, one-way ANOVA comparing the mean values of the mimics (*** P < 0.001).
Figure 10: Functional validation of SUMF1 as a putative target of miR-23a and miR-377 using Western blot and quantitative real-time PCR (qRT-PCR) analyses. Cells were transiently transfected with 50 nM of miR-23a and miR-377 mimics in CHO suspension cells stably expressing a recombinant difficult-to-express (DTE) lysosomal sulfatase protein and cell lysates harvested 48 h post-transfection. (A) Western blot showing SUMF1 protein levels and relative densitometry signal values of SUMF1 protein levels relative to GAPDH (B) Endogenous SUMF1 mRNA levels post-transfection relative to scramble (NT-siRNA) non-targeting control. SUMF1 expression was normalized to GAPDH (C) Endogenous miR-23a and miR-377 levels post-transfection relative to scramble. MiRNA expression was normalized to U6 snoRNA. Anti-SUMF1 siRNA (si-SUMF1) was used as a positive control. Error bars represent the standard error of mean (SEM) of three independent experiments (n=3). For statistical analysis, one-way ANOVA comparing the mean values of the mimics (* P < 0.05; ** P < 0.01; *** P < 0.001).
Figure 11: Transient transfection of 50 nM miR-23a and miR-377 antimiRs in (targeting either endogenous mouse or CHO miR-23a or miR-377) CHO suspension cells stably expressing a recombinant difficult-to-express (DTE) lysosomal sulfatase protein. Cells were cultured in a 6-well plate using a batch process and kept in a static incubator. (A) Viable cell density (VCD) and viability 72 h post transfection (B) Sulfatase titer and (C) specific activity determined by an Elisa and activity assay respectively. Samples were taken from the culture supernatants 72 h post-transfection and data represented as fold change normalized to scramble (NT-siRNA) control. Anti-SUMF1 siRNA (si-SUMF1) was used as a positive control. Error bars represent the standard error of mean (SEM) of three independent experiments (n=3). For statistical analysis, one-way ANOVA comparing the mean values of the (* P < 0.05; ** P < 0.01; *** P < 0.001).
Figure 12: Effect of transient transfection of 50 nM miR-23a and miR-377 antimiRs in (targeting either endogenous mouse or CHO miR-23a or miR-377) CHO suspension cells stably expressing a recombinant difficult-to-express (DTE) lysosomal sulfatase protein using Western blot and quantitative real-time PCR (qRT-PCR) analyses. (A) Western blot of SUMF1 protein levels and relative densitometry signal values of SUMF1 protein levels relative to GAPDH 48 h post transfection. (B) Endogenous SUMF1 mRNA levels post-transfection relative to scramble (NT-siRNA) non-targeting control. SUMF1 expression was normalized to GAPDH. (C) Endogenous miR-23a and miR-377 levels 48 h post-transfection relative to scramble. miRNA expression was normalized to U6 snoRNA. Anti-SUMF1 siRNA (si-SUMF1) was used as a positive control. Error bars represent the standard error of mean (SEM) of three independent experiments (n=3). For statistical analysis, one-way ANOVA comparing the mean values of the (* P < 0.05; ** P < 0.01; *** P < 0.001).
Figure 13: Transient transfection of titrated concentration (10-200 nM) of cgr-miR-23a mimics in CHO suspension cells stably expressing a recombinant difficult-to-express (DTE) lysosomal sulfatase protein. Cells were cultured in a 6-well plate using a batch process and kept in a static incubator. (A) Upper panel: Viable cell density (VCD). Lower Panel: Viability. Both 72 h post-transfection (B) Western blot of SUMF1 protein levels and densitometry signal values of SUMF1 protein levels relative to GAPDH 48 h post transfection (C) Endogenous SUMF1 mRNA levels post-transfection relative to scramble (NT-siRNA) non-targeting control. SUMF1 expression was normalized to GAPDH (D) Endogenous miR-23a levels 48 h post-transfection relative to scramble. miRNA expression was normalized to U6 snoRNA. Anti-SUMF1 siRNA (si-SUMF1) was used as a positive control. Data are mean +/- SEM of two independent experiment. For statistical analysis, unpaired two-tailed t-test was applied (* P < 0.05; ** P < 0.01).
Supporting Information.

Table 16: Primers used for miRNA and mRNA expression analyses using qRT-PCR

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Forward primer (5’ – 3’)</th>
<th>Reverse Primer (5’ – 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAPDH</td>
<td>AGCCACATCGCTCAGACAC</td>
<td>GCCCAATACGACCAAATCC</td>
</tr>
<tr>
<td>U6 snoRNA</td>
<td>CTCGCTTCGGCAGCACA</td>
<td>AACGCTTCAGAAATTTGCGT</td>
</tr>
<tr>
<td>SUMF1</td>
<td>TACCTGTCAAAGGCGCTAAC</td>
<td>CATTCCAGGAGACATGGAGAACC</td>
</tr>
<tr>
<td>cgr-miR-23a-3p</td>
<td>AUCACAUUGCCAGGAUUUCC</td>
<td>Qiagen universal reverse primer</td>
</tr>
<tr>
<td>mmu-miR-23a-3p</td>
<td>AUCACAUUGCCAGGAUUUCC</td>
<td>Qiagen universal reverse primer</td>
</tr>
<tr>
<td>cgr-miR-377-3p</td>
<td>UGAAUCACACAAAGGCAACUUU</td>
<td>Qiagen universal reverse primer</td>
</tr>
<tr>
<td>mmu-miR-377-3p</td>
<td>AUCACACAAAGGCAACUUUUGU</td>
<td>Qiagen universal reverse primer</td>
</tr>
</tbody>
</table>
Supporting Information

Table 17: MiRNAs predicted to target SUMF1 gene in mouse

Prediction tools used: TargetScan (v.7.1), miRanda-mirSVR (Aug’10), Diana-microT-CDS (v5.0), Diana-Tarbase (v.8) in combination with miRecords (April 27, 2013)

miRecords predicted 234 mouse miRNAs which includes 62 mouse miRNAs predicted altogether by the 4 tools to target SUMF1. MiRNAs in “red” were the miRNAs predicted by at least 4 tools and were selected for further downstream analyses. Only data for mir-23a-3p and mir-377-3p is reported.

<table>
<thead>
<tr>
<th>miR</th>
<th>miR-377-3p</th>
<th>miR-6373</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-21</td>
<td>miR-377-3p</td>
<td>miR-6373</td>
</tr>
<tr>
<td>miR-590-5p</td>
<td>miR-882</td>
<td>miR-690</td>
</tr>
<tr>
<td>miR-590-3p</td>
<td>miR-185-5p</td>
<td>miR-743a-3p</td>
</tr>
<tr>
<td>miR-329</td>
<td>miR-504</td>
<td>miR-497b</td>
</tr>
<tr>
<td>miR-362-3p</td>
<td>miR-133a-3p</td>
<td>miR-1896</td>
</tr>
<tr>
<td>miR-181a-5p</td>
<td>miR-133b-3p</td>
<td>miR-350-3p</td>
</tr>
<tr>
<td>miR-181b</td>
<td>miR-133c</td>
<td>miR-7240-5p</td>
</tr>
<tr>
<td>miR-181c</td>
<td>miR-215-3p</td>
<td>miR-350-5p</td>
</tr>
<tr>
<td>miR-181d</td>
<td>miR-7227-3p</td>
<td>miR-291b-3p</td>
</tr>
<tr>
<td><strong>miR-23a-3p</strong></td>
<td>miR-669e-3p</td>
<td>miR-7094-3p</td>
</tr>
<tr>
<td>miR-23b</td>
<td>miR-330-3p</td>
<td>miR-1224-3p</td>
</tr>
<tr>
<td>miR-124-3p</td>
<td>miR-6911-3p</td>
<td>miR-7036b-3p</td>
</tr>
<tr>
<td><strong>miR-106a-5p</strong></td>
<td>not in found CHO</td>
<td>miR-3074-5p</td>
</tr>
<tr>
<td><strong>miR-106b-5p</strong></td>
<td>miR-5625-3p</td>
<td>not in found CHO</td>
</tr>
<tr>
<td><strong>miR-17-5p</strong></td>
<td>miR-875-3p</td>
<td>not in found CHO</td>
</tr>
<tr>
<td><strong>miR-20a</strong></td>
<td>miR-3473a</td>
<td>miR-300-3p</td>
</tr>
<tr>
<td>miR-20b-5p</td>
<td>miR-3473c</td>
<td>miR-204-5p</td>
</tr>
<tr>
<td><strong>miR-93-5p</strong></td>
<td>miR-7678-3p</td>
<td>miR-29-5p</td>
</tr>
<tr>
<td>miR-290-5p</td>
<td>miR-3082-3p</td>
<td>miR-22-3p</td>
</tr>
<tr>
<td>miR-292-5p</td>
<td>miR-3089-3p</td>
<td>miR-381-3p</td>
</tr>
<tr>
<td>miR-3470b</td>
<td>miR-155-5p</td>
<td></td>
</tr>
</tbody>
</table>
**Supplementary Figure**

**Figure 14:** si-hSUMF1 was titrated to determine the optimal concentration that will reduce or completely delete exogenously hSUMF1 stably expressed in CHO cell line secreting a sulfatase: **A)** Western blot analysis of different concentration of si-hSUMF1 and volume of transfection reagent used. **B)** Relative signal intensity of hSUMF1 levels in CHO cell line exogenously expressing hSUMF1 compared to mock, scramble and untransfected cells.